



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

MARCO ANTONIO LASMAR ALMADA

SIMULAÇÃO DE FENÔMENOS SOCIAIS: ASPECTOS EPISTEMOLÓGICOS E
METODOLÓGICOS

Campinas

2018

MARCO ANTONIO LASMAR ALMADA

SIMULAÇÃO DE FENÔMENOS SOCIAIS: ASPECTOS EPISTEMOLÓGICOS E
METODOLÓGICOS

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Engenharia Elétrica, na área de Engenharia de Computação.

Orientador: Prof. Dr. Romis Ribeiro de Faissol Attux

Este exemplar corresponde à versão final da dissertação defendida pelo aluno Marco Antonio Lasmar Almada, e orientada pelo Prof. Dr. Romis Ribeiro de Faissol Attux.

Campinas

2018

Agência(s) de fomento e nº(s) de processo(s): Não se aplica.

ORCID: 0000-0002-0127-6549

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

AL61s Almada, Marco Antônio Lasmar, 1991-
Simulação de fenômenos sociais : aspectos epistemológicos e metodológicos / Marco Antônio Lasmar Almada. – Campinas, SP : [s.n.], 2018.

Orientador: Romis Ribeiro de Faissol Attux.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Interação social - Simulação por computador. 2. Ciências sociais - Metodologia. 3. Engenharia de software. I. Attux, Romis Ribeiro de Faissol, 1978-. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Simulation of social phenomena : epistemological and methodological aspects

Palavras-chave em inglês:

Social interaction - Computer simulation

Social sciences - Methodology

Software Engineering

Área de concentração: Engenharia de Computação

Titulação: Mestre em Engenharia Elétrica

Banca examinadora:

Romis Ribeiro de Faissol Attux [Orientador]

Juliano Souza de Albuquerque Maranhão

Cristiano Cordeiro Cruz

Data de defesa: 30-10-2018

Programa de Pós-Graduação: Engenharia Elétrica

COMISSÃO JULGADORA — DISSERTAÇÃO DE MESTRADO

Candidato: Marco Antonio Lasmar Almada

Data da Defesa: 30 de outubro de 2018

Título da Tese: Simulação de Fenômenos Sociais: Aspectos Epistemológicos e Metodológicos

Prof. Dr. Romis Ribeiro de Faissol Attux (Presidente)

Prof. Dr. Juliano Souza de Albuquerque Maranhão

Prof. Dr. Cristiano Cordeiro Cruz

A ata de defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

À gentileza de estranhos.

Agradecimentos

Em primeiro lugar, gostaria de agradecer ao professor Romis Attux, não só pela enriquecedora experiência da orientação acadêmica, como também pelo pleno apoio na “guinada teórica” que este projeto sofreu desde sua formulação inicial até a versão atual. Mas, para além da dimensão estritamente acadêmica, também devo agradecer a Romis por me ensinar, pelo exemplo, como a especialização do conhecimento pode ser conciliada com os interesses humanísticos — *homo sum; humani nil a me alienum puto* — e com a preocupação com o próximo.

Gostaria também de agradecer aos professores Juliano Maranhão e Cristiano Cruz, por sua disposição em participar de minha banca de qualificação e da banca final. Ao longo do processo de redação desta dissertação, ambos forneceram comentários que enriqueceram significativamente o texto final.

Este projeto, embora só tenha assumido a forma atual há relativamente pouco tempo, tem suas raízes em tópicos que me cativam de longa data. Seria, portanto, omissos se não mencionasse minha gratidão à professora Katia Alves Rabello de Araújo e a Shridhar Jayanthi, que me abriram as portas da vida acadêmica, ao professor Sidney Chalhoub, cujas aulas de Introdução ao Estudo da História me despertaram o interesse por questões metodológicas nas humanidades, e a Diego Vicentin, por meio de quem travei o primeiro contato com a obra de Simondon e os estudos de ciência, tecnologia e sociedade. Destaco estes nomes como representativos dos muitos docentes que me ajudaram a definir o tipo de acadêmico que pretendo vir a ser.

Devo, também, agradecer a conhecidos e amigos que me apoiaram nos planos intelectual e pessoal ao longo de meu mestrado. Qualquer lista que fizesse aqui certamente deixaria de fora alguns nomes importantes, mas ainda assim tentarei destacar alguns nomes. Dentre os amigos da vida universitária, sou particularmente grato a Murilo Favaretto, Alessandro Andrioni, Alan Godoy, Monique Oliveira, João Schwarz, Eric Velten e Luca Giannotti. De outras esferas da vida, devo mencionar o apoio vindo de Renata Vaz Shimbo, Lays Stanziani, João Gabriel Arato Ferreira, Andrea Faggion, Valter Alnis Bezerra, Bárbara Barros Barbosa, Rafael Viana, Mahayana Godoy, Sylvia Pinheiro, Halis Alves, Clarissa Borges, Guilherme Paulino Passos e Melina Fioravante. Os aportes destes interlocutores certamente resultaram em um trabalho mais sólido, cujas lacunas restantes são de minha inteira responsabilidade.

Por último, mas não menos importante, sou grato a meu pai, Marco Antonio, e minha tia Isabel Cristina pelo apoio moral e material ao longo dos anos.

“The critic is not the one who debunks, but the one who assembles. The critic is not the one who lifts the rugs from under the feet of the naïve believers, but the one who offers the participants arenas in which to gather. The critic is not the one who alternates haphazardly between antifetishism and positivism like the drunk iconoclast drawn by Goya, but the one for whom, if something is constructed, then it means it is fragile and thus in great need of care and caution.”
(Bruno Latour)

Resumo

Simulações computacionais são usadas dentro e fora da academia para a modelagem de fenômenos sociais. Ainda assim, há pouca discussão específica sobre a fundamentação teórica do uso de simulações em domínios como o direito, a sociologia ou a administração, dentre outros, uma vez que a literatura acadêmica sobre simulações sociais foca em questões metodológicas de aplicação, o que abre espaço para a rejeição das simulações por pesquisadores menos propensos ao uso de métodos formais em domínios sociais. A partir da análise de casos de uso de simulações sociais, esta dissertação identifica uma definição de simulação compatível com os papéis em que simulações são empregadas. Tal definição, por sua vez, torna possível reconciliar o uso de simulações com os diferentes modos de explicação adotados em ciências sociais, expandindo o escopo de aplicação das técnicas simulacionais. Por fim, as considerações teóricas são aplicadas ao projeto de simulações computacionais, possibilitando a construção de simulações que contribuam para o entendimento e a tomada de decisões nos domínios sociais modelados.

Palavras-chaves: Interação social - Simulação por computador; Ciências sociais - Metodologia; Engenharia de software.

Abstract

Computer simulations are used inside and outside of academia for modeling social phenomena. Still, there is a lack of specific discussion on the theoretical underpinnings of applying simulation to domains such as law, sociology, or management, as the academic literature on social simulation focuses on methodological questions, opening space to the wholesale rejection of simulations by researchers averse to formal methods in social domains. Drawing from the analysis of actual simulations of social phenomena, this thesis identifies a definition of simulation that is compatible with the roles in which such simulations are used. That definition, in turn, allows for the reconciliation between simulations and the many modes of explanation adopted in social science, expanding the scope of application of simulation techniques. Finally, those theoretical remarks are deployed to the project of computer simulations, making it possible to build simulations that contribute to understanding and decision-making in the modeled social domains.

Keywords: Social interaction - Computer simulation; Social sciences - Methodology; Software engineering.

Sumário

0	Introdução	12
1	Simulações computacionais	14
1.1	Aplicações não-acadêmicas de simulações	16
1.2	O que é uma simulação computacional?	18
1.3	Verificação e validação de simulações computacionais	21
1.4	Simulações como ferramenta retórica	23
1.5	Simulações e os fenômenos sociais	26
2	Simulações e conhecimento	28
2.1	Simulações como experimentos in silico	29
2.1.1	O fecho epistêmico das simulações	29
2.1.2	Simulações computacionais e aleatoriedade	30
2.2	Simulações como experimentos mentais	32
2.2.1	Opacidade e entendimento das simulações	34
2.2.2	Simulações como experimentos mentais de larga escala	35
2.3	Conhecimento: uma visão operacional	36
2.3.1	A formulação clássica de conhecimento	38
2.3.2	Respostas ao problema de Gettier	39
2.4	Simulação e representação	42
2.4.1	Simulações, teoria e observações	44
2.5	O status científico das simulações	47
3	Formalização de fenômenos sociais	48
3.1	Simulações e os modos de explicação social	51
3.1.1	Simulações como ferramenta de explicação causal	52
3.1.2	Simulações e as explicações não-causais	54
3.1.3	Explicação e Predição	56
3.2	Papéis analíticos das simulações	57
3.3	Papéis retóricos das simulações	60
3.4	Os usos das simulações	63
4	Desenvolvimento de simulações	64
4.1	Modelagem formal de sistemas	66
4.1.1	Modelos probabilísticos	69
4.1.2	Objetivos de modelagem	70
4.1.3	Complexidade e reducionismo	71
4.1.4	Ficções e idealizações de modelagem	72
4.2	Coleta e uso de dados	73
4.3	Implementação e execução de simulações	75

4.4	Calibração e validação de simulações	77
4.5	Análise e uso dos resultados	80
4.6	Limites não-técnicos das simulações	81
5	Conclusão	83
5.1	Uma imagem das simulações computacionais	85
5.2	Limites e possibilidades das simulações	86
	 Bibliografia	 88

0 Introdução

— *Entretanto, construí na minha mente um modelo de cidade do qual extrair todas as cidades possíveis — disse Kublai. — Ele contém tudo o que vai de acordo com as normas. Uma vez que as cidades que existem se afastam da norma em diferentes graus, basta prever as exceções à regra e calcular as combinações mais prováveis.*

Italo Calvino, *As Cidades Invisíveis*: As Cidades e o Nome 1

Simulações computacionais são sistemas computacionais desenvolvidos para descrever o comportamento de sistemas em cenários nos quais a solução analítica de um modelo matemático que represente o sistema é inviável ou não é interessante. Nas ciências sociais, estes objetos foram empregados por pesquisadores interessados em fenômenos sociológicos, linguísticos e econômicos, dentre outros ([ELSENBROICH; GILBERT, 2013](#)); e, a partir destas aplicações, os métodos de simulação ganharam espaço também em aplicações comerciais e de políticas públicas ([DESAI, 2012](#)), dentre outros fins não-acadêmicos.

Boa parte da discussão teórica a respeito de simulações computacionais é feita a partir de exemplares extraídos da prática das ciências naturais e da engenharia.¹ Ainda que exista uma literatura dedicada às simulações de fenômenos sociais (e.g. [Gilbert e Troitzsch \(2005\)](#)), tais obras em geral focam em apresentação de técnicas computacionais e exemplos de casos de uso. Assim, é escassa a análise sobre temas diversos como as dimensões éticas do uso de simulações² e a adequação das simulações aos diferentes modos de pesquisa empregados nas ciências sociais.

Para sanar tal lacuna, propomos um estudo da prática acadêmica e não-acadêmica das simulações de fenômenos sociais, investigando os modos de uso e as características comuns entre estas ferramentas. Este estudo panorâmico, apresentado no início do Capítulo 1, serve de base para introduzirmos, na Seção 1.2, a formalização da noção de simulação computacional proposta por [Humphreys \(2004\)](#), a partir da qual identificamos os elementos centrais das simulações: o modelo que produz as saídas da simulação, a implementação computacional deste modelo, e a correspondência entre as saídas do modelo e o sistema que se deseja replicar. Exploramos, então, estes elementos para identificar como uma simulação pode ser usada enquanto ferramenta para persuadir terceiros da correteza de análises nela baseadas.

¹ Como exemplos, temos [Weisberg \(2013\)](#) e [Saam \(2017b\)](#).

² Ponto particularmente relevante quando se tem em conta que simulações de sistemas sociais muitas vezes são usadas como ferramentas para determinar intervenções no sistema que elas modelam, o que pode dar origem a injustiças e problemas jurídicos que raramente são relevantes em sistemas físicos.

Definida a noção de simulação computacional de um fenômeno social, abordamos no Capítulo 2 a discussão a respeito de como simulações poderiam gerar conhecimento relativo ao seu objeto de estudo. Como salientam [Frigg e Reiss \(2009\)](#), os problemas relacionados à produção de conhecimento — em especial conhecimento científico — a partir das simulações computacionais são redutíveis a problemas filosóficos mais gerais, mas mesmo esta visão reducionista mantém em aberto a necessidade de identificar as operações adequadas de redução pertinentes. Antes disso, no entanto, sustentamos nas seções 2.1 e 2.2 que simulações devem ser entendidas como uma forma particular de experimento mental, e não como um processo similar aos experimentos científicos. A partir desta visão, discutimos na Seção 2.4 como este tipo de experimento mental pode produzir uma representação do fenômeno estudado capaz de servir de base para conhecimento, nos termos da Seção 2.3.

Uma vez estabelecida a possibilidade em abstrato de se obter conhecimento a respeito de um fenômeno social por meio de simulações, abordamos, no Capítulo 3, como um fenômeno social em específico pode ser formalizado através de uma simulação. Na Seção 3.1, exploramos como simulações podem ser usadas não só para explicar fenômenos sociais sob uma perspectiva de identificação de mecanismos causais, mas também para amparar outros modos de explicação adotados nas ciências sociais, como a identificação de razões para a ação. Em tais explicações, como em outros usos, as simulações desempenham um papel instrumental que discutimos na Seção 3.2; em particular, elas são empregadas como ferramentas de *persuasão*, tanto na prática científica como fora dela, aspecto que é objeto da Seção 3.3.

No Capítulo 4, examinamos o ciclo de desenvolvimento de simulações descrito por [Siegfried \(2014\)](#), aplicando a cada etapa os *insights* teóricos identificados nas discussões anteriores. A partir desta análise, sugerimos práticas de desenho de *software* que podem ser empregadas para assegurar que as simulações sejam usadas de um modo compatível com seus objetivos cognitivos e práticos.

Por fim, sintetizamos no Capítulo 5 as principais conclusões teóricas deste percurso e as recomendações práticas que dele seguem, consolidando o caso em favor do uso de simulações computacionais no estudo de fenômenos sociais. As simulações podem contribuir para o refino de modelos teóricos — ainda que elas não prescindam destes —, ao estabelecer uma linguagem comum para a validação de hipóteses e explicitação de fatores implícitos, facilitando a colaboração em equipes multidisciplinares e permitindo que os resultados das simulações sejam replicáveis de maneira independente. Desta forma, o uso de técnicas de simulação computacional nas linhas aqui descritas pode contribuir para o entendimento de dinâmicas sociais que ocorrem em um mundo cada vez mais complexo ([MILLER; PAGE, 2007](#)) e marcado por riqueza de dados e de sistemas de informação.

1 Simulações computacionais

Kublai era um atento jogador de xadrez; seguindo os gestos de Marco, observava que certas peças implicavam ou excluía a proximidade de outras peças e deslocavam-se de acordo com certas linhas. Transcurando a variedade de formas, ele definia a disposição de um objeto em relação ao outro sobre o pavimento de maiólica. Pensou: “se cada cidade é como uma partida de xadrez, o dia em que eu conhecer as suas regras finalmente possuirei o meu império, apesar de que jamais conseguirei conhecer todas as cidades que este contém.”

Italo Calvino, *As Cidades Invisíveis*: 8

Simulações não são uma técnica estranha ao estudo de fenômenos sociais. Um dos pioneiros na adoção desta ferramenta foi o economista americano Thomas Schelling, que propôs, na década de 1970, um modelo idealizado para descrever as dinâmicas de segregação espacial em uma sociedade: um *grid* retangular é parcialmente povoado com atores¹ de uma dentre duas classes. Caso um ator tenha muitos vizinhos de uma classe distinta da sua, ele optará por mudar sua posição, em busca de uma vizinhança que lhe seja mais aceitável² (ELSENBROICH; GILBERT, 2013, p. 69–70).

O modelo de Schelling foi proposto com finalidades didáticas,³ permitindo a visualização do impacto que leves mudanças nas condições iniciais do modelo podem ter no cenário final. Nesta abordagem de *autômato celular*, os atores possuem um conjunto bem limitado de ações possíveis: podem apenas se mudar ou permanecer na célula que ocupam. Uma abordagem mais sofisticada é a dos *modelos baseados em agentes*: um agente computacional⁴ é um ator interno ao sistema computacional dotado de múltiplas possibilidades de interação com seu ambiente e com outros agentes. Tais modelos encontraram vários usos nas ciências sociais, como estudos de fenômenos sociológicos (BIANCHI;

¹ Usaremos, ao longo da dissertação, o termo *ator* para nos referirmos a qualquer tomador de ações, sejam estas ações internas a um sistema computacional, sejam elas externas — por exemplo, alguém que tome uma decisão a partir dos resultados de uma simulação.

² O modelo de segregação emprega uma variável binária para distinguir entre os atores, normalmente representada por uma cor. Um resultado surpreendente é que, a depender da distribuição inicial da população, mesmo atores que tenham apenas uma leve preferência por conviver com seus semelhantes acabarão se mudando, o que leva a população a se agrupar em *clusters* (ELSENBROICH; GILBERT, 2013).

³ O modelo de segregação foi objeto de várias implementações computacionais (ELSENBROICH; GILBERT, 2013). Curiosamente, o próprio Schelling não era favorável ao uso de computadores na execução de seu modelo; em uma discussão que ele escreveu a respeito do tratamento computacional da segregação, cita que simulações executadas no computador não resolvem o problema da análise dos resultados e podem comprometer o valor didático da simulação ao obscurecer algumas das etapas intermediárias (SCHELLING, 1972), antecipando algumas das discussões que retomaremos neste capítulo.

⁴ Ao longo da dissertação, as referências a *agente* devem ser entendidas como tratando de agentes computacionais.

SQUAZZONI, 2015), da emergência e evolução de gramáticas (STEELS, 2016) e de diversos outros fenômenos coletivos (VENTURINI; JENSEN, P.; LATOUR, 2015).

A partir da década de 1980, as técnicas de simulação em ciências sociais passam a fazer maior uso de recursos computacionais, em especial com o surgimento dos computadores pessoais. Um exemplo paradigmático desta fase é o estudo de Axelrod (1984) a respeito da evolução da cooperação. Como parte do estudo, os autores propuseram uma competição entre algoritmos com diferentes estratégias de cooperação em um dilema do prisioneiro iterado, determinando as estratégias mais aptas a partir dos resultados da competição entre implementações propostas por pesquisadores de diversas instituições.⁵ Aqui, as simulações atuaram como um campo de teste de elementos da teoria dos jogos, mas a relação entre estas áreas é uma via de duas mãos, uma vez que *insights* oriundos da teoria dos jogos podem ser incorporados às técnicas de construção de modelos baseados em agentes (ELSENbroich; GILBERT, 2013, p. 95–97).

Além dos autômatos celulares, dos modelos baseados em agentes e das simulações baseadas em teoria dos jogos, outra aplicação de simulações computacionais em domínios sociais é o uso de processos de geração de dados envolvendo métodos de Monte Carlo,⁶ de forma a obter estimativas mais precisas de estatísticas sobre uma população. Um exemplo é o artigo de Gelman e King (1994) em que os autores estabelecem um modelo de sistema eleitoral bipartidário baseado em distritos e, a partir da geração de várias eleições simuladas por amostragem estatística, descrevem como tal modelo pode ser utilizado para avaliar propriedades do sistema eleitoral, como o viés partidário nas votações.

Abordagens baseadas em simulações computacionais como as descritas acima⁷ são empregadas por cientistas sociais no estudo de objetos diversos. Nos últimos anos, simulações passaram a ser usadas no estudo de fenômenos emergentes em sistemas sociais, em que técnicas computacionais são usadas para verificar o surgimento de comportamentos e instituições a partir de perspectivas evolutivas e de teoria de jogos (CONTE et al., 2012, p. 328–329). Com o crescimento do poder computacional à disposição dos pesquisadores e do volume de dados a que eles podem ter acesso, surge a expectativa na comunidade científica de que abordagens baseadas em *big data* possam dar origem a modelos que descrevam fenômenos mais complexos que os estudados até o momento, como dinâmicas

⁵ A solução vencedora, *tit for tat*, foi um programa de poucas linhas submetido por Anatol Rapoport com uma estratégia simples — repetir a ação feita pelo adversário enfrentado na rodada anterior — que prevaleceu mesmo contra estratégias mais sofisticadas (AXELROD, 1984).

⁶ Métodos de Monte Carlo são simulações que usam números aleatórios para computar o resultado de uma computação que pode ser dispendiosa demais de realizar (BEISBART; NORTON, 2012). Um exemplo clássico é o cálculo do valor de π : em um quadrado de lado 1, inscreve-se um quarto de uma circunferência de raio 1. A partir disso, a simulação gera um grande número de pontos aleatórios dentro do círculo, e a razão entre o número de pontos dentro do quarto de circunferência e o número total de pontos gerados é usada como aproximação de $\frac{\pi}{4}$.

⁷ Mas não necessariamente limitadas às classes de abordagem apresentadas.

culturais (CONTE et al., 2012, p. 337–338) ou a formação de opiniões individuais e coletivas na sociedade (NONNENMACHER et al., 2014).

A partir dos exemplos discutidos, é possível identificar alguns papéis que as simulações computacionais podem desempenhar na investigação de fenômenos estudados pelas ciências sociais. Elas podem fornecer uma forma de *aproximação* de cálculos cuja forma analítica pode ou não ser solucionável dentro das condições de contorno de experimento; neste caso, a simulação funciona mais diretamente como uma forma de extensão da capacidade computacional humana (HUMPHREYS, 2004). Outro uso comum é a *validação* de modelos teóricos já existentes: os resultados da simulação são comparados com padrões já existentes, em busca da melhor explicação para um dado fenômeno. Um terceiro uso possível de simulações é para a *exploração* de teorias: os resultados da simulação são investigados em busca de inconsistências oriundas do modelo subjacente ou de interações inicialmente não previstas. Estes e outros usos, melhor discutidos no Capítulo 2 do presente texto, compartilham um papel *cognitivo*, uma vez que se encaixam na atividade de construção e refinamento do conhecimento científico.

1.1 Aplicações não-acadêmicas de simulações

Simulações computacionais de fenômenos sociais também são empregadas em contextos que não a produção de conhecimento científico, em especial por empresas interessadas em prever o comportamento de consumidores e outros atores sociais em busca de ferramentas para guiar sua tomada de decisão em cenários com incertezas. Ao simular cenários possíveis de mercado e avaliar os impactos esperados em cada situação, uma empresa⁸ pode ajustar seu curso de ação em busca dos resultados desejados. Técnicas simulacionais foram empregadas por empresas em vários contextos em que um tratamento analítico seria difícil ou pouco adequado, como o uso de modelos de agentes para descrever alterações em modelos de negócios (LIEDER; ASIF; RASHID, 2017), de métodos de Monte Carlo em aplicações de engenharia financeira (GLASSERMAN, 2003) ou de abordagens de fluxos para modelar a entrada de pacientes na emergência de hospitais britânicos (MOHIUDDIN et al., 2017).⁹

Simulações computacionais são empregadas por agentes de mercado como uma forma de entender o comportamento de fenômenos, desempenhando aqui papéis cognitivos semelhantes aos descritos no uso acadêmico das simulações, como a descrição do fator humano em modelos de linhas de produção (PAIVA, 2010). Todavia, há uma ênfase nas pos-

⁸ Ou outro ator inserido em um contexto social, como desenhistas de políticas públicas.

⁹ Este caso, em particular, mostra como o uso de simulações fora da academia não está restrito a necessidades estritamente mercadológicas, uma vez que o sistema de saúde britânico, ainda que inserido em uma economia de mercado, segue um modelo público de financiamento. A discussão desta seção, em linhas gerais, continua aplicável a modelos cuja preocupação central não é mercadológica, mas a abundância de aplicações comerciais motivou o foco nestas.

sibilidades de tomada de ação motivadas pelos resultados das simulações, que atuam como uma fonte de dados para os processos de valoração e extração de valor dos consumidores descritos por [Fourcade e Healy \(2016, p. 16, tradução própria\)](#): “Conforme traços digitais de comportamentos individuais são agregados, armazenados e analisados, mercados veem pessoas a partir de uma lente de merecimento e não-merecimento.”¹⁰ Nestes cenários, a simulação computacional não necessariamente desempenha um papel cognitivo, mas sim é voltada para produzir outros efeitos em seus usuários e em terceiros.

Toda simulação computacional é construída a partir de modelos conceituais do sistema que pretende representar. A descrição abstrata fornecida por esses modelos é concretizada a partir de sua implementação em dispositivos computacionais realmente existentes, ainda que estes ocasionalmente sejam ocultados por abstrações, e sua criação é, em geral, motivada pelo interesse de pessoas ou instituições interessadas em prever comportamentos sociais. O impacto de uma simulação, no entanto, não está restrito às pessoas que consomem os dados, uma vez que decisões tomadas a partir de modelos analíticos podem ter impacto em *stakeholders* que estão distantes dos resultados da simulação, na medida em que estes produzem consequências em suas vidas, como a concessão ou não de crédito bancário.¹¹

O uso das simulações pode, assim, afetar as escolhas disponíveis a humanos direta ou indiretamente ligados ao objeto simulacional. Essa *mediação* das interações entre pessoas e o mundo pode afetar de diversas formas as interações entre indivíduos ou de um indivíduo com o ambiente em que ele está inserido. Uma primeira forma pode ser entendida como o estabelecimento de uma *relação hermenêutica* entre o indivíduo e a simulação ([ROSENBERGER; VERBEEK, 2015, p. 16](#)), na qual a atenção humana é direcionada para a interpretação que a tecnologia possibilita sobre o mundo. No exemplo do empréstimo bancário, a simulação é orientada ao processo de decisão de um gerente humano, que interpretará a situação financeira do cliente com base nos resultados computacionais.

Este cliente, por sua vez, dificilmente terá qualquer forma de contato com os resultados da simulação. Ainda assim, existe, entre o cliente e a simulação, uma *relação de fundo*,¹² uma vez que as saídas da simulação afetam indiretamente as possibilidades de ação deste cliente diante do seu ambiente — por exemplo, ao negar ou fornecer o capital que o cliente precisa para abrir seu próprio negócio.¹³ Nestas relações, a presença de

¹⁰ No original: “As digital traces of individual behaviors are aggregated, stored, and analyzed, markets see people through a lens of deserving and undeservingness”.

¹¹ [Fourcade e Healy \(2016\)](#) propõem um *framework* teórico para entender como organizações atuantes no mercado tendem a coletar grandes volumes de dados a respeito de indivíduos e usar estes dados para medir e classificar as ações individuais, estabelecendo uma *economia de juízo moral* em que os acontecimentos são vistos como merecidos ou não com base nas ações e gostos anteriores dos indivíduos classificados.

¹² No original, *background relation* ([ROSENBERGER; VERBEEK, 2015, p. 18](#)).

¹³ As relações hermenêuticas e as relações de fundo entre humanos e a tecnologia — no nosso caso, as simulações computacionais — surgem na obra de Don Ihde, que também propôs duas dinâmicas

uma simulação (praticamente) não é notada na operação normal, e só se torna relevante em casos de falha evidente: se a simulação passa a ser objeto de um processo judicial por adotar um algoritmo discriminatório, sua existência deixa de ser algo ignorado pelas pessoas afetadas. Mas, de forma geral, o efeito de uma simulação usada na tomada de ação não pode ser entendido apenas a partir do estudo de seus impactos nos consumidores imediatos dos dados.

O uso voltado ao mercado das simulações computacionais ajuda a explicitar o papel das simulações como *mediadoras* da interação entre indivíduos e o mundo. Este papel está presente nas simulações empregadas em contexto acadêmico, uma vez que as transformações epistêmicas geradas pela análise dos resultados de uma simulação “acadêmica” também mudam a forma como o objeto de estudo é percebido, alterando os modos de interação do indivíduo com seu ambiente. As diferenças de finalidade de uso podem levar a uma ênfase no papel cognitivo das simulações ou no papel de mediação, mas ambos são elementos constituintes das simulações computacionais.

1.2 O que é uma simulação computacional?

Simulações são utilizadas de diversas formas na investigação de fenômenos sociais, como é visível nos exemplos mercadológicos e acadêmicos anteriores, a partir dos quais podemos identificar algumas características compartilhadas entre os vários objetos que chamamos de simulações computacionais. Esta seção buscará delimitar quais são esses elementos comuns para, a partir disso, definir a classe das simulações computacionais que será objeto desta dissertação.

Todos os exemplos de simulações anteriormente tratados buscam, de uma forma ou de outra, reproduzir o comportamento de um sistema externo, seja ele passível de representação por uma equação que tentamos resolver de forma aproximada ou um contexto social que não conseguimos modelar de forma explícita. Com base nisso, uma primeira tentativa de definição seria:

Definição 1.1 *Uma simulação é uma reprodução do comportamento de um sistema-alvo ao longo do tempo.*

Nos casos descritos pela Definição 1.1 e que nos interessam nesta dissertação, a simulação é realizada através da implementação computacional de um sistema lógico

adicionais. Nas relações de *alteridade* (ROSENBERGER; VERBEEK, 2015, p. 18), a pessoa interage com a simulação de uma forma interativa, similar à interação que ele teria com outras pessoas; no caso das simulações, esta dinâmica poderia ser verificada tanto em cenários de realidade aumentada quanto em jogos como Civilization (ALMADA; ATTUX, 2018, p. 19). Já uma relação de incorporação — no original, *embodiment* — pode ser entendida como a transformação das capacidades de um indivíduo perceber o mundo e atuar nele (ROSENBERGER; VERBEEK, 2015, p. 14), por exemplo, ao ter sua visão melhorada por um par de óculos. Nesta dissertação, focaremos nas relações de fundo e hermenêuticas.

que pretende descrever um sistema externo, que serve de alvo simulacional. A reprodução é verificada por meio da comparação dos resultados do programa executado com as observações feitas do sistema-alvo. Incorporando estes elementos, temos uma definição construtiva de simulação computacional:

Definição 1.2 *Uma simulação computacional é um sistema computacional que reproduz o comportamento de um sistema-alvo ao longo do tempo.*

A Definição 1.2, em uma primeira passagem, captura a noção intuitiva de simulações computacionais formada a partir dos exemplos já apresentados, funcionando como uma regra de reconhecimento que permite, por exemplo, diferenciar uma simulação de uma resolução analítica de um modelo matemático. Todavia, a construção realizada deixa uma grande margem de ambiguidade nos termos empregados, que analisaremos a seguir.

Um primeiro ponto que merece atenção é a expressão “sistema computacional”. Todos os exemplos vistos trabalham com simulações implementadas em computadores reais, ainda que fazendo uso de abstrações computacionais como a “nuvem”. No entanto, é possível, ao menos em tese, que algoritmos capazes de reproduzir o comportamento de um sistema-alvo em uma máquina de Turing não possam ser implementados em um sistema concreto, seja por escassez de recursos computacionais, seja porque o tempo necessário para execução exceda os limites toleráveis para a aplicação concreta da simulação. Temos, então, um cenário em que a reprodução é teoricamente possível mas não ocorre na prática.

A noção de reprodução de comportamento também merece mais atenção. Em primeiro lugar, é necessário entender que tipo de comportamento a simulação deve buscar reproduzir: basta gerar saídas que repliquem as observações do sistema-alvo ou também se visa reproduzir o estado interno deste? Uma vez determinado o tipo de reprodução pertinente, é necessário definir quais são os critérios para comparar o comportamento dos dois sistemas. Também é possível questionar o caráter temporal da reprodução: uma simulação precisa produzir os resultados de maneira sequencial e na mesma ordem que o sistema-alvo, ou é suficiente que a simulação gere saídas que correspondam individualmente a estados possíveis do sistema-alvo?

Uma tentativa de solucionar as questões identificadas no conceito de simulação pode ser vista em [Humphreys \(2004\)](#). Para este autor, uma *simulação-núcleo*¹⁴ é um sistema S que produz saídas que são soluções de um modelo que representa corretamente o sistema-alvo B . A modelagem em si pode ser feita através de modelos matemáticos de equações diferenciais resolvidas de maneira aproximada, mas nada impede que a re-

¹⁴ No original, *core simulation* ([HUMPHREYS, 2004](#), p. 121).

apresentação seja feita a partir de um modelo estocástico,¹⁵ das propriedades emergentes do sistema, ou qualquer outra técnica de modelagem que possa ser reproduzida em um computador. Como define [Humphreys \(2004, p. 121, tradução própria\)](#):¹⁶

Definição 1.3 *Um sistema S fornece uma simulação-núcleo de um objeto ou processo B apenas se S é um dispositivo computacional concreto que produz, através de um processo temporal, soluções para um modelo computacional (...) que representa B corretamente, de maneira estática ou dinâmica. Se, além disso, o modelo computacional usado por S representa corretamente a estrutura do sistema real R , então S fornece uma simulação-núcleo do sistema R em relação a B .*

Aqui, temos que a representação é necessariamente computacional, o que já restringe o escopo para as modalidades de simulação que abordaremos nesta dissertação. A representação pode ser *estática* — isto é, com uma saída invariante no tempo — ou *dinâmica*. Simulações estáticas podem ser úteis na descrição de sistemas sociais ou como resultados intermediários, mas na maior parte dos casos fazemos uso de simulações dinâmicas, que tentam capturar os aspectos temporais dos eventos sociais modelados no dispositivo computacional ([ELSENbroich; Gilbert, 2013](#)).

A definição apresentada acima dá margem para simulações que capturem não somente a percepção externa que se tem do objeto ou processo modelado, mas também possam capturar a própria estrutura do sistema real. A saída deste tipo de simulação consiste na união entre o estado externo observável e uma representação do estado interno do sistema-alvo, mas não apresenta diferenças conceituais em relação à saída produzida por uma simulação que enfatize apenas o aspecto externo. O desacoplamento destas duas categorias, portanto, é mais relevante para a modelagem em si do que para a discussão a respeito do que é uma simulação computacional, já que a modelagem das estruturas pode ser representada como a união entre um modelo que representa o processo e um modelo que representa as observações feitas do sistema-alvo.

Da mesma forma, a questão da *corretude* da simulação estará ligada à modelagem adotada: a representação de um sistema discreto envolverá um mapeamento entre as saídas da simulação e os estados observados no sistema-alvo, enquanto uma simulação

¹⁵ Isto é, uma descrição quantitativa de um fenômeno que prevê um conjunto de resultados possíveis e as probabilidades de cada um desses cenários ([Pinsky; Karlin, 2010, p. 1](#)). Por exemplo, o lançamento de um dado honesto pode levar a seis cenários, cada um com a mesma probabilidade, em que um número diferente cai na face superior do dado.

¹⁶ No original:

System S provides a core simulation of an object or process B just in case S is a concrete computational device that produces, via a temporal process, solutions to a computational model (...) that correctly represents B , either dynamically or statically. If in addition the computational model used by S correctly represents the structure of the real system R , then S provides a core simulation of system R with respect to B .

de saídas contínuas recorrerá a alguma forma de *aproximação*, em especial no caso de fenômenos estocásticos. Embora a ideia de representação científica seja assunto de uma série de debates acadêmicos (GHINS, 2016), o uso de simulações não introduz, a princípio, nenhuma questão nova ou mudança nos termos em que este debate é posto.

Outra distinção introduzida pela definição, desta vez de maneira implícita, é a separação entre a simulação-núcleo e uma simulação completa. Para Humphreys (2004), uma simulação-núcleo se torna uma simulação completa quando acrescida de uma *representação* que ordene as *saídas* do núcleo de forma a capturar corretamente o comportamento do sistema-alvo ao longo do tempo. As formas como as saídas de uma simulação podem representar observações de um sistema-alvo serão assunto discutido ao longo desta dissertação. Porém, acompanhando Frigg e Reiss (2009), discordamos que a escolha específica de representação das saídas seja um elemento *essencial* do objeto simulacional: o que interessa para que uma simulação descreva adequadamente um fenômeno é a representação R_1 que o modelo interno à simulação faz do sistema simulado. Uma vez identificada uma R_1 satisfatória, seria possível, ao menos em princípio, construir várias formas R_2 de representar as saídas produzidas a partir do modelo fundado em R_1 para o usuário final — por exemplo, através de tabelas ou animações visuais — que seriam igualmente válidas. Desta forma, nos interessa a noção de simulação-núcleo adotada por Humphreys (2004), que adotaremos como definição de simulação computacional.¹⁷

Por fim, a definição operacional de Humphreys (2004) é que a simulação computacional possui um componente *concreto*: um sistema computacional concreto é necessário para fornecer uma simulação do sistema-alvo. A simulação não é só um modelo matemático que descreve um sistema: ela está realizada em um sistema computacional específico e sujeita às limitações deste.¹⁸ Mais do que isso: a simulação é sempre executada em máquinas realmente existentes, o que limita os recursos computacionais efetivamente disponíveis, o tempo necessário para a execução do programa que a implementa, dentre outras limitações práticas a uma execução completa da simulação e à validação dos resultados obtidos.

1.3 Verificação e validação de simulações computacionais

Um primeiro desafio na construção de simulações computacionais de sistemas externos está na verificação da correteza dos processos computacionais utilizados para

¹⁷ O estudo em separado de R_2 pode ser interessante, por exemplo, para melhor entender as visualizações de dados que podem contribuir para os papéis que uma simulação desempenha, tema que não é abordado nesta dissertação.

¹⁸ No paradigma vigente de computabilidade, que tem como abstração fundante as máquinas descritas em Turing (1937), podem ser encontrados limites de ordem prática, como a existência de problemas para os quais não se conhece algoritmo capaz de fornecer soluções em tempo polinomial, bem como limitações à própria capacidade da máquina abstrata, como a impossibilidade de uma solução geral para o problema da parada.

gerar as soluções do modelo empregado para representar o sistema-alvo. Processos de verificação são empregados ao longo do ciclo de desenvolvimento de *software* para assegurar que o programa em construção atende os requisitos estabelecidos para seu aceite.¹⁹

Erros de computação podem aparecer em diversos pontos da execução de uma simulação. Muldoon (2007) apresenta uma tipologia de fontes de erros computacionais. A classificação parte de um nível baixo de abstração: o próprio *hardware* em que a simulação é executada pode alterar o resultado das operações. Um exemplo trazido por Muldoon (2007) é o *bug* FDIV das primeiras gerações de processadores Pentium, que introduzia erros constantes nas operações de ponto flutuante. Além destes casos em que o comportamento do *hardware* é diferente do comportamento esperado, também existe a possibilidade de defeitos no equipamento causados por falhas de fabricação ou por desgaste ao longo da vida útil; por exemplo, a corrupção de dados em um disco rígido por dano físico ao dispositivo. Por fim, erros de *hardware* podem ser causados por fatores externos ao dispositivo computacional, como oscilações no fornecimento de energia.

Em um nível maior de abstração, Muldoon (2007) identifica uma segunda classe de erros, agora oriundos do uso de linguagens de programação e bibliotecas pré-existentes. Alguns destes erros estão presentes nas próprias ferramentas empregadas. No caso de ferramentas mais estabelecidas, em que o processo de desenvolvimento solucionou problemas identificados nas versões iniciais, uma fonte de erro mais comum é o desconhecimento, por parte dos desenvolvedores de uma simulação, de detalhes de implementação das ferramentas empregadas, como métodos de arredondamento. Em particular, a falta de informação em relação às soluções utilizadas pelos desenvolvedores das ferramentas empregadas pode levar a dificuldades para diagnosticar problemas em casos de uso pouco frequentes ou com documentação desatualizada ou inexistente.

A terceira categoria de erros abordada em Muldoon (2007) surge durante a implementação da simulação propriamente dita. Tais erros podem surgir de uma conversão inadequada do modelo em alto nível para um algoritmo implementável, ou por falha humana durante a implementação do algoritmo. Esta etapa é aquela sobre a qual os desenvolvedores da simulação possuem maior controle, podendo empregar ferramentas e boas práticas de desenvolvimento para minimizar os erros.

Os três níveis de erros descritos por Muldoon (2007) podem resultar em programas cujas saídas não são condizentes com os resultados esperados com base no modelo subjacente à simulação, ou mesmo levar o programa a não concluir sua execução, seja pela ocorrência de um erro fatal, seja pela introdução de um laço infinito no programa. Assim, a verificação exaustiva de qualquer modelo não-trivial exigiria a avaliação de um grande

¹⁹ Sommerville (2011, p. 207, traduções próprias) conceitua verificação ao discutir testes de *software*: “O objetivo da verificação é conferir se o *software* atinge seus requisitos funcionais e não-funcionais declarados”; já a validação é um processo distinto que tem por objetivo “(...) garantir que o *software* atinja as expectativas do cliente.”

número de instruções computacionais, inviabilizando a verificação manual das operações.

A validade de uma simulação não deve ser confundida com sua verificação; enquanto esta assegura que o código implementado corresponde à especificação proposta, aquela se preocupa em garantir que as saídas da simulação representam corretamente o sistema que se pretende modelar. Para alguns modelos, a validação das saídas é uma avaliação da coerência interna do modelo, o que pode introduzir vieses de confirmação: a simulação é avaliada com base em sua aderência aos modelos teóricos, mas ao mesmo tempo as saídas da simulação fornecem respaldo a tais modelos (VENTURINI; JENSEN, P.; LATOUR, 2015).

A capacidade de generalização de uma simulação é interessante para seu uso na investigação de fenômenos, seja porque o acesso aos dados empíricos é caro demais (VENTURINI; JENSEN, P.; LATOUR, 2015, p. 1.3), seja porque o modelo será usado para prever resultados que ainda não podem ser observados diretamente — como a projeção de vendas de um produto ainda não lançado e que apresenta certas diferenças em relação ao portfólio existente. Nestes casos, a simulação é uma ferramenta disponível para que o pesquisador possa “examinar as consequências lógicas das presunções teóricas feitas no nível local” (VENTURINI; JENSEN, P.; LATOUR, 2015, 1.3, tradução própria),²⁰ avaliando a coerência interna de seus modelos teóricos em busca de contradições. Mesmo que o modelo sobreviva a este teste de coerência interna, a validade das inferências feitas com base nos modelos treinados a partir de dados escassos pode se mostrar limitada — por exemplo, se os dados amostrados não forem suficientemente representativos do fenômeno —, possibilidade que exige tratamento especial na validação das saídas da simulação. No Capítulo 4 da presente dissertação, discutiremos abordagens para a detecção, prevenção e mitigação destes níveis de falha, cuja ocorrência impede que a simulação produza resultados aproveitáveis para as análises das ciências sociais.

1.4 Simulações como ferramenta retórica

Uma simulação computacional de um fenômeno social é construída com a pretensão de representar, a partir do modelo a ela subjacente, o comportamento de indivíduos ou grupos sociais em uma determinada situação. Técnicas de verificação e validação como as descritas na seção anterior contribuem para fortalecer a pretensão de que os resultados produzidos da simulação correspondem, de alguma forma, ao fenômeno descrito, o que estabeleceria a validade das inferências feitas a partir dos resultados de simulação mesmo quando estas se referem a casos distintos daqueles observados empiricamente.

O consumidor *imediato* da simulação — a(s) pessoa(s) ou sistema(s) que possuem acesso direto aos resultados das execuções da simulação — poderá analisar as saídas

²⁰ No original: “examine the logical consequences of the theoretical assumptions made at the local level”

geradas através de ferramentas estatísticas, lógicas ou teóricas, dentre outros.²¹ As inferências feitas a partir desta análise podem servir de substrato para a discussão da própria simulação e dos modelos que a originaram, ou podem também ser empregadas em argumentações voltadas a consumidores *mediatos* da simulação, cuja interação com as saídas computacionais dar-se-á de maneira indireta. Por exemplo, em um projeto de lei sobre segurança pública que incluía um estudo de criminologia computacional na sua fundamentação, os legisladores dificilmente interagirão com as simulações de forma direta, mas ainda assim a sua decisão de votar contra ou a favor do projeto — para não falar no impacto que este, uma vez aprovado, tenha na sociedade — será reflexo, em parte, de análises influenciadas por uma simulação. Os dois usos das inferências extraídas de simulações envolvem papéis linguísticos distintos, que serão melhor explorados na Seção 3.3, mas compartilham um ponto central: em ambos os casos, a simulação é empregada como fonte de legitimidade para os argumentos dela derivados, desempenhando papel *retórico*.

Salientar que as simulações executam um papel retórico não significa, em si, remover delas o caráter científico: como aponta McCloskey (1998, cap. 1), o discurso científico em si é uma forma de escrita voltada à persuasão, cujo alvo primário é a comunidade científica da área de estudo. Sob esta perspectiva, as simulações e seus modelos subjacentes atuam como *metáforas não-ornamentais*, que estabelecem analogias entre as variáveis de simulação e os comportamentos a elas prescritos nos modelos, bem como enfatizam as dimensões de comparação capturadas pela simulação (MCCLOSKEY, 1998, cap. 3).

Argumentos que fazem uso de simulações computacionais são, assim, vulneráveis a críticas à validade destas. A Seção 1.3 apresentou abordagens que podem ser usadas para a verificação e validação dos resultados de simulações computacionais, mas tais abordagens são baseadas no entendimento que os criadores da simulação têm do sistema-alvo. Uma primeira ameaça a essa validade, portanto, está na qualidade de tal entendimento: um modelo que represente o sistema-alvo de uma maneira incompatível com o entendimento científico vigente na área de estudos dificilmente será aceito como argumento favorável a uma tese.

Um exemplo, apontado por Venturini, Pablo Jensen e Latour (2015), pode ser visto no uso de modelos baseados em agentes para integrar modelos de micro-comportamento com macrodescrições em um nível mais alto de abstração, “reconstru[indo] empiricamente a continuidade da existência coletiva.”²² Para os autores do artigo, as abordagens típicas de modelos baseados em agentes não resolvem este problema, já que pressupõem que estes não podem atuar no fenômeno de maneira global (VENTURINI; JENSEN, P.; LATOUR, 2015, 1.6, tradução própria):

²¹ O Capítulo 4 traz uma discussão mais extensa sobre a análise dos resultados de simulações.

²² No original: “reconstruct[ing] empirically the continuity of collective existence”

[Na “tragédia dos comuns”], interesses pessoais levam atores a usar em excesso um bem compartilhado (um pasto comum, por exemplo) em detrimento da comunidade. Supondo a existência de agentes atômicos, cada um com seu interesse individualmente definido (como exige a teoria dos jogos), a maior parte desta literatura não pode senão confirmar a exploração excessiva do bem comum. Contudo, trabalho empírico conduzido por Elinor Ostrom (1990) mostrou que cooperação humana pode (...) encontrar arranjos para a gestão bem-sucedida dos comuns. Simulações sociais falham em obter estes arranjos, pois elas desconsideram os mecanismos sutis que governam o estabelecimento da confiança necessária para cooperação.

A existência de tais falhas, sob esta perspectiva, impediria que as simulações pudessem ser tomadas como uma representação fidedigna dos sistemas sociais estudados, invalidando os argumentos baseados em simulações que não consigam dar origem aos arranjos possíveis para tratar os comuns. Todavia, não existe nenhum obstáculo *a priori* que impeça a construção de uma simulação que incorpore “mecanismos sutis”²³ como os descritos, de forma que a invalidade aqui descrita parece ser mais um problema de técnica simulacional do que uma invalidação geral das simulações computacionais neste problema.

O argumento de [Venturini, Pablo Jensen e Latour \(2015\)](#) tangencia uma crítica comum às simulações: a ideia de que a técnica em si seria ideologicamente carregada, noção que encontra raízes em uma tradição mais ampla de crítica ao uso de modelos matemáticos na descrição de fenômenos sociais. Em parte, esta linha surge diante do uso de abordagens típicas da ciência econômica — dentre elas o recurso a modelos formais — no estudo de questões tipicamente associadas à história, à ciência política e outras ciências sociais, movimento caracterizado por [McCloskey \(1998\)](#) como um imperialismo da Economia que teria em Gary Becker uma figura de destaque. Becker, ganhador do Nobel, é um dos principais nomes da chamada *Escola de Chicago*, associada a uma abordagem economicamente liberal cujos pressupostos são refletidos nos modelos empregados.²⁴

A escolha das premissas empregadas em um modelo pode introduzir significativo viés político ou metodológico em um sistema, mesmo que seus construtores não percebam. Neste sentido, é esclarecedora a análise que [Wolff \(2013\)](#) faz dos modelos formais empregados por John Rawls e Robert Nozick em suas célebres obras no terreno da filosofia política. Ambos os autores, segundo [Wolff \(2013\)](#), fazem uso de uma linguagem

²³ A validade de uma simulação depende da representação adequada do sistema-alvo, mas esta não necessariamente envolve a representação exata de cada um dos fatores envolvidos: “Apesar de [ser] ocasionalmente exigida e superficialmente atraente, a nuance inibe a abstração de que a boa teoria depende.” ([HEALY, 2017](#), p. 121, tradução própria). Sob esta perspectiva, a abstração seria uma característica de um bom modelo teórico e, por extensão, de uma simulação computacional.

²⁴ [McCloskey \(1998\)](#) traz como exemplo a aproximação que Becker faz ao considerar uma família como equivalente a uma firma de pequenas dimensões, empregando, portanto, as teorias econômicas da firma para descrever a atividade familiar.

carregada de terminologia da teoria dos jogos, associando, aos pressupostos desta, suas próprias premissas. Tais pontos de partida adicionais é que, no fundo, são responsáveis pelos resultados *derivados* dos métodos formais empregados.

É prematuro, todavia, associar qualquer uso de modelos matemáticos a uma posição política “neoliberal”: em [Roemer \(1982\)](#), temos o uso de modelos de teoria dos jogos e da teoria da escolha racional para o estudo das noções de exploração e classes sociais sob uma perspectiva marxiana, informada por teóricos como G. A. Cohen. Este trabalho, inserido no contexto do chamado *marxismo analítico*, serviu de ponto de partida para análises quantitativas informadas por teorias diversas das adotadas no *mainstream* econômico. Em particular, temos em [Cogliano e Jiang \(2016\)](#) uma sistematização de propostas para o uso de modelos baseados em agentes tanto a partir de perspectivas econômicas pós-keynesianas quanto em um programa de pesquisa mais próximo do marxismo clássico.

Simulações não precisam ser ideologicamente neutras para ter algum valor científico. A crítica a essa pretensão de neutralidade é útil para refinar a construção dos modelos e situá-los propriamente em meio ao ferramental teórico-metodológico ao dispor de um especialista de ciências sociais, mas não inviabiliza o uso de simulações por parte do pesquisador adepto de qualquer vertente ideológica ou metodológica, já que este não só pode construir simulações que reflitam seus pressupostos como também submeter à crítica simulações construídas por terceiros, bem como interpretar os resultados produzidos por estas simulações a partir de um prisma ideológico diverso do adotado pelo construtor original da simulação. Simulações podem ter — e, na prática, terão — vieses ideológicos, mas isso é um reflexo das concepções prévias de seus construtores, não uma propriedade das técnicas em si.

1.5 Simulações e os fenômenos sociais

Simulações computacionais são construídas com base em modelos dos fenômenos que desejam descrever. No entanto, o caráter distintivo destas simulações surge a partir de sua implementação em dispositivos computacionais concretos que tornam viável o uso de simulações para abordar problemas em que uma solução analítica é inviável ou indesejável, ao mesmo tempo que introduzem modos de falha que não estariam presentes em um modelo puramente lógico, como a possibilidade de falhas de *hardware*.

A crescente disponibilidade de recursos computacionais levou à adoção de diversas técnicas de simulação aplicadas a múltiplos domínios das ciências sociais, com graus variados de aceitação por parte do *mainstream* de cada ciência. O uso científico das simulações computacionais também serviu de ponto de partida para a construção de simulações voltadas a problemas aplicados, como os que surgem em contextos comerciais. Em

ambos os casos, as simulações desempenham tanto um papel cognitivo no entendimento dos fenômenos quanto um papel operacional, ao servir de ponto de partida para ações no mundo externo ao modelo; a predominância de cada papel depende do caso concreto.

Uma simulação concreta deve produzir saídas que representem, de alguma forma, os estados de seu sistema-alvo. Para isso, é necessário que o programa que a implementa execute de maneira correta as operações previstas, e que se estabeleça uma analogia entre as saídas do programa e as observações do fenômeno. Neste sentido, técnicas de engenharia de *software* podem assegurar, nos limites já discutidos, a corretude das operações, mas não bastam para vincular adequadamente o resultado de uma simulação às observações. Nos próximos capítulos, identificaremos as condições para que uma simulação justifique as pretensões de rigor e validade em que se baseia seu uso na argumentação científica ou baseada em aplicações científicas.

2 Simulações e conhecimento

Você sabe melhor que ninguém, sábio Kublai, que jamais se deve confundir uma cidade com o discurso que a descreve. Contudo existe uma ligação entre eles.

Italo Calvino, *As Cidades Invisíveis*: As Cidades e os Símbolos 5

O uso de simulações no estudo de fenômenos sociais tem como pressuposto a ideia de que é possível aprender algo sobre um dado objeto de estudo a partir da análise dos resultados de uma simulação que o represente. Uma vez estabelecida uma relação entre as saídas computacionais e o fenômeno de interesse, as simulações podem ser empregadas para prever estados futuros, para validar modelos explicativos e para avaliar cenários contrafactuais, dentre outros usos.

Mas o que torna válidas as inferências feitas com base em simulações computacionais? O componente prático desta questão — isto é, como construir uma simulação específica que seja válida para uma questão particular sob investigação — será discutido no Capítulo 4, a partir das técnicas de engenharia de *software* e pesquisa social. Antes de discutir as abordagens cabíveis no caso particular, no entanto, é necessário entender como as simulações computacionais se relacionam com os objetos que pretendem descrever.

Há, na literatura, quatro conjuntos principais de posições sobre o tema: as que consideram a simulação uma espécie de experimento *in silico* e aquelas que percebem-na como uma forma de experimento mental, além das teses de que as simulações diferem em essência de cada um dos dois modelos anteriores (SAAM, 2017b, p. 295). Sustentaremos aqui que a visão de simulações enquanto um tipo de experimento mental marcado pela opacidade de seus processos internos é a melhor explicação para o papel que as simulações computacionais desempenham na prática científica, em virtude da relação destes objetos com os sistemas que pretendem descrever.

Estabelecida esta natureza metodológica, procederemos a uma análise do papel da simulação na formação do conhecimento científico. De um lado, as simulações podem atuar na validação interna dos modelos teóricos empregados em sua construção, nos termos do Capítulo 1. De outro, elas permitem a descoberta de conhecimento novo através da *explicitação* de consequências não-óbvias ou até contraintuitivas de seus modelos subjacentes. A partir destes dois aspectos, é possível delinear como as simulações podem servir de base para a justificação de proposições a respeito dos sistemas que pretendem descrever.

2.1 Simulações como experimentos in silico

A expressão “experimento *in silico*” tem origem na pesquisa em ciências biológicas, na qual é empregada para designar a reprodução de processos naturais (ou feitos em laboratório) através de simulações computacionais (MIRAMONTES, 1992). O termo busca estabelecer paralelos com modalidades de experimentação, como os experimentos *in vitro*,¹ difundidas em campos como a biologia, sustentando que os processos reproduzidos em computador podem ser explorados de forma análoga à exploração de sistemas biológicos reais. Esta visão encontra eco em alguns filósofos da ciência, para quem as simulações estariam mais próximas do experimento de fato do que de um experimento mental, já que as conclusões baseadas em simulações não seriam produzidas por dedução e sim pela inferência a partir dos resultados computacionais (SAAM, 2017b).

O processo de experimentação pode ser entendido, na formulação de Hacking, como uma forma de criar, produzir, refinar e estabilizar fenômenos (SAAM, 2017b, p. 300). Neste cenário, buscamos entender como o comportamento de uma *variável independente* muda quando provocamos alterações controladas no comportamento de uma ou mais *variáveis dependentes*. Esta busca por controle está presente, também, nas simulações computacionais, nas quais os desenvolvedores conseguem controlar os valores de parâmetros simulados e as regras lógicas do modelo subjacente. Para além deste nível procedimental, também é possível afirmar que as simulações produzem resultados similares aos dados gerados a partir de experimentos (WINSBERG, 2009) ou mesmo que elas funcionam como *instrumentos de medida* para o comportamento do sistema-alvo (MORRISON, 2009). Em todos estes casos, a interação com os resultados simulacionais seria um processo de descoberta de conhecimento que diria algo sobre o sistema-alvo da simulação.

2.1.1 O fecho epistêmico das simulações

De forma geral, tratar as simulações como experimentos significa dizer que, a partir das análises dos resultados computacionais, é possível fazer inferências que não são consequências lógicas automáticas das premissas de implementação do modelo. Um mecanismo para isso, descrito por van der Grient (2011, p. 18), pode ser identificado na ideia de que simulações poderiam ser usadas para superar o *fecho epistêmico*.² Nas palavras de Medeiros (2013, p. 10), “O ‘fecho epistêmico’ é o princípio que afirma que o conhecimento é fechado sob implicação”; ou seja, um agente que conhece uma proposição *p* conheceria todas as proposições que seguem de *p* por implicação lógica.

¹ Experimentos realizados com seres vivos, mas sem que estes estejam em seu contexto habitual; por exemplo, a replicação de um microorganismo em uma placa de Petri.

² Em inglês, *closure under known implication*. Como aponta Medeiros (2013), os princípios de fecho epistêmico ainda são uma questão em aberto na epistemologia, e devem ser analisados de acordo com o tipo de problema epistêmico e com os agentes modelados.

A geração de resultados de simulações, como discutido na Seção 1.3, envolve uma sequência de passos computacionais cuja verificação individual é impraticável, tornando *opaco* o funcionamento da simulação (DI PAOLO; NOBLE; BULLOCK, 2000). Como resultado, a produção dos resultados — e, por extensão, as inferências neles baseados — envolve operações que não são logicamente evidentes e, por isso, poderiam resultar em proposições além das contidas no fecho epistêmico do sistema lógico original, já que o resultado das computações, por não ser evidente, não faria parte do conjunto inicial de fatos (VAN DER GRIENT, 2011, p. 18).

Com base no parágrafo anterior, temos um mecanismo através do qual as simulações podem produzir conhecimento novo, mas encontra dois obstáculos. O primeiro deles está na dependência da opacidade: embora a dificuldade de verificar os passos individuais da inferência dificulte a visualização de todas as operações lógicas envolvidas, os processos de verificação de *software* podem, em princípio, estabelecer que os passos tomados são consequências lógicas necessárias de seus antecedentes.³ Mesmo que a opacidade seja suficiente para superar o fecho, um segundo obstáculo surge na relação com o sistema-alvo: a implicação lógica não é suficiente para garantir que as proposições novas também se apliquem ao sistema descrito, já que elas podem acabar revelando pontos de divergência entre o sistema e o modelo que tenta representá-lo. Então, passa a ser necessário encontrar uma forma de validar tais inferências, o que esvazia a pretensão de que a simulação seja suficiente para, a partir da superação do fecho, descobrir informações sobre o sistema-alvo que não estejam contidas no modelo inicial.⁴ A argumentação de van der Grient (2011) não é, portanto, suficiente para estabelecer a simulação como fonte de conhecimento novo sobre o sistema estudado, mas será retomada mais adiante para situar o papel das simulações na produção de conhecimento.

2.1.2 Simulações computacionais e aleatoriedade

Outra possível justificativa para que as simulações computacionais consigam agregar conhecimento além do presente em seu modelo lógico é a *aleatoriedade*. Simulações computacionais podem empregar, em sua execução, fontes de números aleatórios, de forma que seu comportamento não poderia ser previsto a partir de uma inspeção do modelo anterior à execução deste. Assim, a investigação das saídas de uma simulação seria a única forma de entender o comportamento do modelo e, na medida em que este se mostra válido, traria informações novas sobre o sistema-alvo da investigação.

Para Beisbart e Norton (2012, p. 411-412), a introdução de fatores aleatórios em uma simulação de Monte Carlo pode até contribuir para que os resultados se assemelhem a resultados experimentais, mas tal semelhança seria apenas superficial, já que a

³ Estas provas de corretude dos operadores lógicos podem ser inviáveis na prática, mas a mera possibilidade de aplicação de métodos formais como os discutidos na Seção 1.3 é suficiente aqui.

⁴ A validação, contudo, pode ser útil para a análise exploratória dos modelos (GELFERT, 2016).

argumentação baseada em simulações é, no fundo, uma inferência a partir das premissas de implementação e não do estudo das saídas por ela produzidas. Estes autores sustentam tal posição com base em dois argumentos. Em primeiro lugar, as simulações não interagem com o sistema-alvo durante sua execução; assim, não possuem nenhuma fonte de informação a respeito do comportamento desse sistema que não esteja contida em suas premissas. Por mais que a fonte de aleatoriedade traga informações externas à formulação lógica simulação, estas novas informações não carregam conhecimento sobre o sistema a ser estudado.⁵

O segundo ponto levantado por [Beisbart e Norton \(2012, p. 412–415\)](#) é mais específico das simulações de Monte Carlo. Por construção, este tipo de simulação minimiza o papel da aleatoriedade nos resultados, que são aproximadamente estáveis em repetições do experimento. Combinada ao argumento anterior, esta propriedade é suficiente para mostrar que as simulações de Monte Carlo são melhor entendidas como formas argumentativas e não como experimentos de fato, já que não há fonte nova de informação a respeito do sistema-alvo para além das premissas iniciais.

Este segundo argumento sobre a aleatoriedade nas simulações de Monte Carlo é naturalmente estendido para algumas outras formas de simulação, como os algoritmos de Las Vegas.⁶ Porém, precisamos ter algum cuidado antes de aplicá-lo a qualquer simulação que faça uso de números aleatórios, já que é possível, ao menos em princípio, que um determinado arranjo simulacional não busque essa estabilidade nos resultados, mas sim a variação em cada execução.

De forma mais geral, é possível expressar a saídas de uma simulação como um feixe de argumentos condicionados aos valores aleatórios gerados. Por exemplo, o gerador de números aleatórios pode ser usado em uma simulação S para descobrir se um agente A tomará uma dentre três ações possíveis em um determinado instante. Nesse caso, não sabemos de antemão se o agente, no instante t , fará a escolha $A_{t,1}$, $A_{t,2}$ ou $A_{t,3}$, mas sabemos que o resultado será uma destas três ações e que os acontecimentos internos à simulação nos instantes $\tau > t$ tomarão a escolha feita no instante t como um fato prévio a respeito do seu universo.

O cenário descrito é demasiado simples diante das possibilidades reais, já que o número de argumentos condicionais possíveis a partir de uma simulação crescerá com o tempo de execução e com o número de pontos em que a aleatoriedade é relevante, o que impediria a avaliação manual de todos estes argumentos. Ainda assim, a resolução destes

⁵ Esta informação torna desnecessária a distinção entre aleatoriedade verdadeira e pseudoaleatoriedade para fins desta análise, já que nenhuma delas carrega em si informação sobre o sistema-alvo ([BEISBART; NORTON, 2012](#)).

⁶ Num algoritmo de Monte Carlo, a execução termina após um número de iterações, e o resultado tem uma probabilidade de estar incorreto. Já os algoritmos de Las Vegas garantem a correteza do resultado, mas não oferecem garantias quanto ao fim da execução.

argumentos é condicionada à fonte de aleatoriedade e não ao sistema-alvo, de forma que, mesmo que o modelo computacional M seja insuficiente para saber de antemão qual será o resultado final de uma execução da simulação, as saídas da simulação podem ser obtidas por inferência a partir do modelo computacional formado ao estender M com a adição dos valores $\xi_1, \xi_2, \dots, \xi_n$ atribuídos aos números aleatórios usados no modelo para aquela execução. Uma vez que a escolha destes números não depende, em geral, das propriedades observadas do sistema-alvo, a escolha de um dos argumentos que a simulação expressa em detrimento dos outros não traz informações novas a respeito do sistema descrito, mas sim dos processos usados para gerar os números aleatórios.

Diante da natureza argumentativa das simulações de Monte Carlo, há duas possibilidades: ou classificamos as simulações de duas formas — distinguindo entre simulações que podem ser percebidas como experimentos e as demais,⁷ — ou classificamos todas as simulações computacionais em alguma outra categoria metodológica. Como o primeiro ponto de Beisbart e Norton não se restringe aos modelos de Monte Carlo, e identificamos substratos para tentar generalizar o segundo argumento, parece-nos mais sensato buscar uma classificação alternativa para as simulações.

Como alternativa à visão de simulações como experimentos, Beisbart introduz a noção de simulações como um modelos de um experimento possível, em que os comportamentos simulados tentam reproduzir os comportamentos esperados em um experimento sem que, no entanto, cheguem a constituir um experimento de fato (SAAM, 2017b, p. 303). Do ponto de vista metodológico, esta proposta aborda a simulação não como uma prática experimental, mas sim como uma abordagem teórica de estudos de um fenômeno, aproximando-se da próxima abordagem que trataremos: a noção de simulações como experimentos mentais.

2.2 Simulações como experimentos mentais

Experimentos mentais⁸ são uma ferramenta comum no raciocínio científico e no ensino da ciência. Um exemplo clássico é o experimento mental empregado por Galileu para mostrar que a velocidade de queda de um corpo não depende de sua massa. Gendler (2000, p. 40–41, tradução própria) apresenta o experimento da seguinte forma:

Imagine que um corpo leve e um corpo pesado são amarrados um ao outro

⁷ Saam (2017a) propõe uma distinção entre duas categorias de simulações sociais: S_{TE} , que seriam simulações com características epistemológicas e metodológicas típicas dos experimentos mentais, e S_E , que seriam similares a experimentos. Em face da discussão aqui posta sobre a relação das simulações com seus referentes, tal distinção é melhor entendida como uma diferença gradual entre modos de uso e não uma diferença *essencial* entre as categorias, interpretação compatível com a proposta da autora (SAAM, 2017a).

⁸ Também conhecidos como *experimentos de pensamento*, a partir do termo alemão *Gedankenexperiment*, cunhado por Ernst Mach (1838–1916) (SORENSEN, 1992, p. 51).

e lançados de uma altura significativa. Qual seria, para um aristotélico, a velocidade natural esperada dessa combinação? Por um lado, o corpo leve deveria retardar o corpo pesado enquanto o corpo pesado acelera o mais leve, de forma que a combinação de ambos deveria cair com uma velocidade que esteja entre as velocidades naturais de seus componentes. [...] Por outro lado, já que o peso dos dois corpos combinados é maior que o peso do corpo pesado por si só, a combinação deveria cair com uma velocidade natural maior que a do corpo pesado. [...] Mas então a previsão é de que o corpo combinado caia mais rápido e mais devagar que o corpo pesado sozinho. [...] A saída deste paradoxo é supor que a velocidade natural com que um corpo cai independe de seu peso

Gendler (2000, p. 2) define o experimento mental em termos de “casos excepcionais”, que seriam situações em que uma teoria descreve a situação de mais de uma forma, sendo uma delas privilegiada pela teoria; na maior parte dos casos, a escolha entre múltiplas descrições possíveis é resolvida através de uma característica que atua como representante da descrição privilegiada, mas, no caso excepcional, esta característica não está presente ou é insuficiente para resolver a questão, como ocorre no exemplo acima quando amarramos os dois corpos. O experimento mental contribuiria para o conhecimento a respeito de um fenômeno ao forçar a explicação de casos excepcionais (GENDLER, 2000, p. 12), como foi o caso no experimento de Galileu: a inconsistência entre os dois raciocínios para a velocidade de queda implicou que a velocidade de queda de um corpo independe de sua massa. De forma relacionada, Sorensen (1992, p. 3) conceitua os experimentos mentais como casos limítrofes de experimentos, cujo funcionamento se aproxima das histórias curtas (SORENSEN, 1992, p. 6) criadas a partir das proposições que delimitam o cenário.

A visão de experimentos mentais como casos limítrofes é contestada por Cooper (SAAM, 2017b) por não conseguir descrever experimentos mentais que não estejam em forma proposicional. Sob esta perspectiva crítica, uma visão mais adequada seria que o experimento mental fornece um modelo para a construção de modelos de mundos possíveis, cujas condições são delimitadas pelo experimento, o que tornaria uma simulação um procedimento computacional que fornece um *padrão* para a construção de mundos possíveis, representados por diferentes iterações do programa simulacional.

Sejam os experimentos mentais casos limítrofes de experimentação, sejam eles um modelo de mundos possíveis, seus resultados são condicionados pelas consequências lógicas da visão de mundo de quem elaborou o experimento mental ou de quem tenta reproduzi-lo (SAAM, 2017b, p. 295–296).⁹ Por este funcionamento baseado na implicação

⁹ O que, por sua vez, abre margem para uma crítica ao uso desta abordagem em domínios sociais: a ideia de que este condicionamento levaria a resultados etnograficamente limitados, uma vez que as

lógica, o fecho epistêmico, caso válido, será aplicável para os experimentos mentais. Assim, o papel destes experimentos não seria revelar proposições novas, mas sim identificar paradoxos que revelem inconsistências entre crenças científicas (SORENSEN, 1992, p. 122–131). O ganho em conhecimento a partir dos experimentos mentais seria, nesta perspectiva, fruto da eliminação das inconsistências de uma teoria científica.

O tratamento conceitual dos experimentos mentais pode ser usado como substrato para o enquadramento metodológico das simulações computacionais, uma vez que seus resultados podem ser empregados na validação dos sistemas lógicos que expressam teorias sobre um fenômeno. Um tratamento que segue esta abordagem pode ser visto em Di Paolo, Noble e Bullock (2000), para quem simulações computacionais seriam um tipo específico de experimento mental marcado pela *opacidade*. A noção de modelos opacos, descrita na Subseção 2.1.1, está relacionada à impossibilidade de visualizar de maneira clara a totalidade dos processos computacionais que produzem os resultados de uma simulação não-trivial, mesmo que com o auxílio de ferramentas automatizadas.

A opacidade de uma simulação significa que não é possível verificar diretamente o processo de produção de seus resultados. Todavia, não segue disso que o uso de simulações obscurece a descrição de um sistema. Como destaca Humphreys (2004, p. 148), a perda de observação sobre detalhes internos pode, na verdade, contribuir para o entendimento de um sistema, ao abstrair etapas de computação que não possuem importância intrínseca. O sistema seria, nesses casos, melhor entendido em um nível de abstração maior do que o das operações individuais, e a opacidade tornaria o sistema mais inteligível.

Humphreys (2004, p. 149) também aponta para a existência de um segundo tipo de opacidade: aquela derivada de uma complexidade irreduzível do sistema. Há casos em que é impossível ou impraticável formar um modelo que descreva o sistema de forma mais simples que a própria simulação. Exemplos de simulações afetadas por opacidade deste tipo seriam o uso de métodos de Monte Carlo para a resolução de integrais intratáveis por sua alta dimensionalidade, bem como os modelos baseados em agentes. Nestas situações, o uso de uma simulação opaca acaba por melhorar o entendimento do fenômeno estudado ao fornecer uma representação mais simples que as alternativas.

2.2.1 Opacidade e entendimento das simulações

A modelagem do problema também é uma forma pela qual simulações, mesmo com sua opacidade, podem resultar em ganhos no entendimento a respeito de um fenômeno. A prática científica em períodos de normalidade, propõe Kuhn (2012, p. 40–42), é feita em meio a uma rede de compromissos — de natureza conceitual, teórica, instrumen-

premissas envolvidas na construção dos experimentos mentais seriam reflexo de intuições culturalmente circunscritas sem a pretensão de universalidade que costumam ter. Não abordaremos este argumento na dissertação de maneira direta, mas vale salientar a discussão que Monteiro Brito Júnior (2018) faz desta suposta limitação no campo da filosofia do direito.

tal, metodológica e sobretudo metafísica — que determinam a atividade dos cientistas. Muitos desses compromissos não são identificados na forma de regras explícitas, mas sim como *conhecimento tácito*, que estaria ligado à experiência individual. Este tipo de conhecimento pode se referir a um *know-how* de natureza técnica ou, em uma perspectiva cognitiva, aos modelos mentais que o indivíduo desenvolve sobre o mundo [Gavira \(2003, p. 20\)](#), e influencia a interpretação que o experimentador faz de um experimento mental, ao delimitar as possibilidades lógicas de resultado.

Como as simulações computacionais não podem fazer uso direto do conhecimento tácito presente na mente do experimentador, sua implementação exigiria que estas premissas fossem traduzidas em expressões explícitas no modelo subjacente à simulação ou em sua implementação.¹⁰ Este processo de implementação acabaria por aumentar a *penetrabilidade cognitiva* das simulações — isto é, o quanto elas podem ser inspecionadas pelo experimentador —, tornando-as mais passíveis de exploração que outras formas de experimento mental ([DI PAOLO; NOBLE; BULLOCK, 2000](#), p. 6). Ao expressar variáveis relevantes na forma de parâmetros e definir regras explícitas que representam as interações esperadas entre objetos da simulação, o processo de construção de uma simulação dá margem para que estes fatores, normalmente enraizados nos processos cognitivos do usuário, possam ter seu funcionamento alterado. Com isso, aumenta-se o alcance das hipóteses alternativas testáveis, através da flexibilização dos mundos possíveis que podem ser gerados pelo procedimento simulacional.

O aumento da penetrabilidade cognitiva encontra limites na medida em que as tarefas de programação de uma simulação são realizadas com o uso de bibliotecas pré-existentes, com premissas e detalhes de código desconhecidos pelo desenvolvedor da simulação — e por vezes inacessíveis para ele. Mesmo assim, a implementação de uma simulação, ainda que fazendo extenso uso de peças já prontas, terá parâmetros e interações que devem ser definidas por seu criador — e, depois, pelos usuários —, explicitando parte do conhecimento implícito envolvido e permitindo a realização de análises de sensibilidade através da alteração destes fatores trazidos à tona pelo processo de desenvolvimento.

2.2.2 Simulações como experimentos mentais de larga escala

Outro ponto em que as simulações computacionais podem estender as capacidades cognitivas de seus usuários é na escala dos experimentos realizáveis. Experimentos mentais tradicionais costumam envolver um número reduzido de agentes e interações entre

¹⁰ Uma vez que o conhecimento tácito escapa à enunciação direta, o processo de captura não será realizado pela expressão direta deste conhecimento, mas sim de maneira indireta. Como os construtores empregarão seu conhecimento tácito nos processos de desenho, programação e calibração da simulação, descritos no Capítulo 4, o sistema computacional resultante terá embutido em si uma fração do conhecimento tácito — por exemplo, na escolha de valores ótimos para parâmetros —, ainda que não seja possível precisar um bloco de código exato em que tal contribuição ocorre.

eles, uma vez que a cognição humana logo esbarra nos limites da capacidade de manter múltiplos objetos na mente ao mesmo tempo.¹¹

O limite no rastreamento de objetos afeta as atividades de desenvolvimento de uma simulação, restringindo o número de objetos que um desenvolvedor consegue manter na mente; todavia, é possível dividir o esforço de desenvolvimento em vários blocos de trabalho, empregando técnicas de desenvolvimento de *software* como as abordadas no Capítulo 4. Uma vez implementados os elementos, a execução da simulação pode envolver grandes números de objetos, idênticos ou distintos, dentro dos limites do poder computacional à disposição; este limite pode ser expandido através de técnicas de escalabilidade horizontal ou vertical.¹²

As simulações computacionais compartilham propriedades essenciais dos experimentos mentais, uma vez que sua atuação é delimitada pela lógica do modelo a ela subjacente. Assim como os experimentos mentais, seu *modus operandi* consiste na formulação de mundos alternativos possíveis, que são então analisados em busca de inconsistências com o corpo de conhecimento já existente sobre o fenômeno. As diferenças entre as duas abordagens surgem em virtude da opacidade das simulações, e podem trazer ganhos e perdas sobre a inteligibilidade dos modelos. Não são, no entanto, suficientes para estabelecer uma diferença de gênero entre as abordagens,¹³ o que sugere que as simulações, como os experimentos mentais, estariam presas ao fecho epistêmico.

2.3 Conhecimento: uma visão operacional

O processo de construção das simulações busca, através dos processos internos do programa, estabelecer uma relação entre as saídas da simulação e os estados observados do sistema-alvo. Com base em tal relação, a análise dos resultados computacionais poderia gerar conhecimento novo sobre o comportamento do sistema representado, colaborando para as abordagens científicas de fenômenos observados. Todavia, o enquadramento das

¹¹ A complexidade do pensamento humano tem como um de seus fatores limitantes a limitação da capacidade da memória de trabalho humana. A capacidade da memória de trabalho está ligada ao tipo de problema com que se trabalha, e não possui uma explicação universalmente aceita pela literatura, embora a evidência atual pareça mais consistente com a teoria de que a interferência mútua entre representações mentais seja o limitante da memória de trabalho (OBERAUER et al., 2016).

¹² *Escalabilidade horizontal* se refere à habilidade de distribuir a carga computacional e os dados usados por um programa entre várias máquinas, como uma forma de variar a capacidade computacional disponível de acordo com a demanda, o que é possível, por exemplo, através de *frameworks* de computação distribuída. Outra abordagem para o mesmo problema é a *escalabilidade vertical*, que consiste em variar os recursos computacionais disponíveis dentro de um dado sistema — por exemplo, adicionando mais processadores ou RAM ao servidor que executará uma simulação (CATTELL, 2011, p. 13).

¹³ Uma divergência neste sentido pode ser identificada em Lenhard (2018), para quem a transparência é condição necessária dos experimentos mentais. Tendo em vista a natureza iterativa dos processos computacionais, seria contraditório falar em experimento mental opaco. Mesmo levando em conta esta ressalva semântica, sustentamos aqui que a semelhança do modo argumentativo simulacional ainda permitiria classificar os *Gedankenexperimente* e as simulações computacionais como espécies de um mesmo gênero metodológico.

simulações como uma forma de experimento mental sugere fortemente que elas estão sujeitas ao fecho epistêmico, diante das questões levantadas no fim da Seção 2.1. Nestas circunstâncias, é possível obter conhecimento *novo* a partir das simulações?¹⁴

A resposta desta pergunta exige uma análise mais detalhada do que entendemos por conhecimento. Tendo em vista que este assunto é estudado pela epistemologia, uma das áreas tradicionais da filosofia,¹⁵ não é nossa pretensão apresentar um panorama do estado da arte da teoria do conhecimento. O objetivo desta seção é mais modesto: explorar as discussões de uma perspectiva instrumental, para delimitar como as simulações podem contribuir com a produção de conhecimento, em particular na descoberta de fatos que não estejam imediatamente acessíveis ao pesquisador.¹⁶

Em uma perspectiva computacional, o conhecimento seria a relação entre um agente conhecedor e uma proposição,¹⁷ que ele pode considerar¹⁸ verdadeira ou falsa (BRACHMAN; LEVESQUE, 2004, p. 2). O conjunto das proposições conhecidas por um indivíduo simulado é representada em uma *base de conhecimento*, que pode ser atualizada com o tempo, bem como conter conhecimento inicial (RUSSELL; NORVIG, 2010, p. 235). Esta base de conhecimento influenciará as interações internas à simulação, de forma que a modelagem de conhecimento é relevante para simulações baseadas em agentes.

Sob esta perspectiva, o conhecimento à disposição de um agente simulado pode ser distinto daquele disponível para o desenvolvedor da simulação: o agente pode não ter acesso a certos conhecimentos, tomar como verdadeira uma proposição que o

¹⁴ Nesta seção, discutiremos o conhecimento científico, que, para Gavira (2003), “caracteriza-se pelo acolhimento metodológico e sistemático dos fatos da realidade sensível”. Tal concepção parece-nos suficientemente ampla para encampar a prática das ciências sociais, mesmo em correntes metodológicas que evitam a importação de métodos típicos das ciências naturais.

¹⁵ Sosa (2017, p. vii) localiza a origem da teoria do conhecimento moderna nos escritos epistemológicos de Descartes, cuja problemática teria por sua vez raízes no ceticismo de Pirro. Já o tratamento clássico do conhecimento como crença justificada e verdadeira pode ser rastreado até Platão, em *Teeteto* e *Mênon*.

¹⁶ Elgin (2017) faz escolha diversa da que tomamos aqui: em vez de tentar reconciliar o papel das simulações com uma leitura da noção de conhecimento, a autora propõe que esta é insuficiente para explicar aspectos da prática científica — dentre eles o uso de simulações. Para Elgin (2017, p. 12–14), o padrão de sucesso aplicado à prática científica não deveria ser a produção de conhecimento — uma vez que a ciência sempre seria deficitária por este ponto de vista —, mas sim uma concepção de *entendimento* com critérios menos rígidos que os discutidos pela literatura em epistemologia e apresentados nesta seção. Tendo em vista a distância entre as escolhas feitas, não discutiremos esta posição de maneira mais extensa nesta dissertação, mas esta perspectiva pode ser relevante para o leitor interessado no tratamento epistemológico das simulações.

¹⁷ Os tratamentos típicos sobre conhecimento, em especial na computação, costumam considerar proposições como o objeto válido do conhecimento, mas existem exceções à regra. A literatura de simulações computacionais possui diversos exemplos de modelagens de fenômenos normativos, em que é importante a representação computacional de regras (ELSEN BROICH; GILBERT, 2013). De um ponto de vista mais teórico, Fischer (2017) apresenta uma proposta para a representação de conhecimento sobre proposições modais, envolvendo proposições *possíveis* ou *necessárias*, tema sobre o qual Nozick (1981, p. 187–188) já havia ensaiado considerações iniciais.

¹⁸ A relação entre o agente e a sentença pode ser descrita por vários verbos — *A sabe* que p_1 , *B espera* que p_2 — mas Brachman e Levesque (2004, p. 2–3) descrevem estas diferenças como *atitudes proposicionais* que seriam secundárias diante do valor lógico atribuído à proposição.

desenvolvedor sabe ser conjectural, ou mesmo “conhecer” uma premissa falsa, inserida de maneira deliberada ou não no sistema. Assim, o conhecimento possuído pelos agentes internos à simulação não é necessariamente igual ao conhecimento que se pode extrair da análise dos resultados da simulação, mesmo que eles possam refinar suas crenças ao longo da execução do programa.

2.3.1 A formulação clássica de conhecimento

A abordagem computacional de conhecimento, por muito tempo, estabeleceu uma equivalência entre conhecimento e crença, tratando os dois indistintamente como relações equivalentes armazenadas nas estruturas de dados de um sistema (CARNOTA; RODRÍGUEZ, 2010, p. 10). Todavia, a formulação filosófica clássica da noção de conhecimento, conhecida como JTB (*justified true belief*),¹⁹ faz distinção entre as duas noções (GETTIER, 1963, p. 121, tradução própria):

Definição 2.1 (JTB) *S sabe que P se e somente se:*

- (i) *P é verdadeira*
- (ii) *S crê que P*
- (iii) *S tem justificativa para crer em P*

O conhecimento detido por um agente seria, assim, apenas uma parte — na prática, um subconjunto próprio — de seu conjunto de crenças. Embora ambos os tipos de proposição possam ser empregados na orientação das ações de um agente, diferenças qualitativas entre a crença e o conhecimento podem ter impacto em problemas práticos: mudanças nas circunstâncias de uma simulação podem remover a justificativa de uma proposição conhecida, ou mesmo torná-la falsa, sem que o agente deixe de acreditar em sua validade e agir com base nela. Nesses casos, a conduta do agente não possui respaldo externo e, portanto, não se traduz em conhecimento a respeito do mundo.

Como exemplo, uma simulação de uma mola que produza resultados com base somente na Lei de Hooke fornecerá informação incorreta quando a mola estiver fora de sua região elástica, ainda que as saídas sejam condizentes com as observações feitas na região elástica. Se um indivíduo baseia seu conhecimento da mola nessa simulação, cuja correteza na região elástica é conhecida, ele tem justificativa para acreditar que o resultado da simulação é igual ao comportamento do sistema-alvo, mas a proposição não é verdadeira na região inelástica.

Outra possibilidade relacionada é que um agente tenha uma crença que seja verdadeira nas condições da simulação, e que ele tenha uma justificativa para sustentar

¹⁹ Em português, crença justificada e verdadeira.

esta crença, mas que a veracidade não seja consequência da justificativa. Cenários deste tipo são conhecidos como *casos de Gettier*, e mostram que as condições acima não são suficientes para estabelecer que o agente *S* de fato *sabe* que *P*. [Gettier \(1963, p. 122, tradução própria\)](#) propõe um exemplo elucidativo:

Suponha que Smith e Jones se candidataram para um certo emprego. Suponha também que Smith tem fortes evidências para a seguinte proposição conjuntiva:

(d) Jones é quem conseguirá o emprego, e Jones tem dez moedas no bolso.

A evidência que Smith tem para (d) pode ser que o presidente da companhia lhe assegurou que Jones seria escolhido no final das contas e que ele, Smith, tinha contado as moedas no bolso de Jones dez minutos atrás. Da proposição (d), segue:

(e) A pessoa que conseguirá o emprego tem dez moedas em seu bolso.

Suponhamos que Smith veja que (d) acarreta em (e), e aceite (e) com base em (d), para a qual ele tem fortes evidências. Neste caso, Smith está claramente justificado em acreditar que (e) é verdadeira.

Mas imagine, além disso, que, sem que ele saiba, o próprio Smith, e não Jones, conseguirá o emprego. E, também sem que Smith saiba, ele mesmo tem dez moedas no bolso. A proposição (e) é, então, verdadeira, ainda que a proposição (d), da qual Smith inferiu (e), seja falsa.

Cenários deste tipo podem ocorrer tanto com os agentes internos a uma simulação quanto na interpretação dos resultados computacionais. Em ambos os casos, as inferências realizadas pelo agente com base em suas crenças justificadas se mostram verdadeiras, mas por motivos que não estão ligados à justificação putativa. Nesse caso, fica difícil sustentar que o agente de fato *saiba* a proposição em que crê, ainda que estejam caracterizados os três elementos da definição de JTB.

2.3.2 Respostas ao problema de Gettier

Ao mostrar que a fórmula “conhecimento é crença justificada e verdadeira” não era suficiente, o artigo de Gettier deu origem a diversas tentativas de definição de conhecimento. Alguns autores optam por adicionar critérios adicionais à formulação JTB; um exemplo é a proposta de *rastreamento da verdade* ([NOZICK, 1981, p. 179, tradução própria](#)):

Definição 2.2 (Rastreamento da verdade) *Um agente S conhece a proposição p através do método M se e somente se:*

- (i) *p é verdadeira.*
- (ii) *S acredita, através do método M , que p .*
- (iii) *Se p não fosse verdadeira e S usasse M para determinar se p , então S não acreditaria, através do método M , que p .*
- (iv) *Se p fosse verdadeira e S usasse M para determinar se p , então S acreditaria, através do método M , que p .*

As condições de rastreamento da verdade tentam condicionar o método de justificação empregado — por exemplo, a proposição (d) do exemplo de Gettier acima não se encaixaria nas condições postas —, mas podem encontrar dificuldades em cenários como a justificação através da união de métodos individualmente insuficientes para este fim, bem como para a identificação de quais métodos são empregados e quais prevalecem sobre outros. Esta falta de resolução, reconhecida pelo próprio [Nozick \(1981\)](#), pode afetar cenários em que um agente simulado age com base em múltiplas regras, bem como a interpretação de *ensembles* de simulações, nos quais o resultado final do programa é um agregado de resultados produzidos por componentes individuais.

A proposta de Nozick é um exemplo do que [Greco \(2015, p. 424\)](#) descreve como uma tentativa de resolver o problema de Gettier através da identificação de equilíbrios coerentes entre as intuições que temos sobre os contornos gerais da imagem de conhecimento e nossas intuições sobre o conhecimento em casos particulares. [Greco \(2015, p. 426–427\)](#) destaca como limitações destes métodos a propensão a gerar construções *ad hoc*, que acomodem as intuições sem fornecer explicações, e também a ausência de critérios para decisão entre modelos que expliquem as intuições de forma igualmente boa ou ruim.

Em face destas limitações explicativas das abordagens baseadas no equilíbrio reflexivo, e da dificuldade em encontrar outras soluções persuasivas para o problema,²⁰ discussões contemporâneas sobre conhecimento tentam explicar o valor do conhecimento em contraste com as crenças verdadeiras, bem como o papel do conhecimento em guiar a comunicação e a ação humanas. Tendo estes pontos em mente, [Greco \(2015, p. 431–432, tradução própria\)](#) propõe que o conceito de conhecimento deve estar ligado a seu valor prático:

[...] o conceito de conhecimento serve para governar o fluxo de informação acionável em uma comunidade epistêmica. Isto é, o conceito governa tanto a

²⁰ Para [Floridi \(2004\)](#), não é possível resolver o problema de Gettier através da modificação das condições do JTB, já que a coordenação entre verdade e justificação seria equivalente a um problema insolúvel em lógica epistêmica.

absorção quando a distribuição de informação, ao serviço da ação individual e da coordenação de ação em grupo.

Esta delimitação da noção de conhecimento nos permite, mesmo sem nos comprometermos com uma ou outra definição explícita, identificar um papel para as simulações no quadro geral do conhecimento científico. Os resultados produzidos pelas simulações estão sujeitos ao fecho epistêmico das teorias usadas em sua construção, mas a análise dessas saídas pode explicitar consequências do modelo subjacente que, de outra forma, escapariam à atenção dos responsáveis pela simulação. Neste papel, as simulações fornecem pontos de partida para investigações subseqüentes ([GELFERT, 2016](#), p. 84) e explicações em potencial para fenômenos ([GELFERT, 2016](#), p. 87).

Como exemplo do uso de simulações na exploração cognitiva, temos um modelo baseado em agentes sobre a cultura Anasazi nos Estados Unidos, que explicitou, através da comparação entre dados arqueológicos e os resultados das simulações, a combinação entre degradação ambiental e pressões socioculturais que levou essa população a abandonar o vale Long House ([GUMERMAN; DEAN, 2000](#)). A execução de várias iterações do programa permitiu aos cientistas analisar a importância relativa dos fatores que a teoria empregada considerava relevantes, repetindo dinâmicas históricas a partir das condições iniciais de uma forma que não é tipicamente associada às disciplinas históricas.

Desta forma, a contribuição das simulações computacionais para o conhecimento viria na medida em que elas desempenham papéis como elementos dos fluxos de distribuição de informação em uma comunidade epistêmica, como veremos na discussão no Capítulo 3 a respeito dos usos retóricos das simulações, na medida em que a análise de seus resultados é usada como lastro científico para discursos relativos a fenômenos sociais. Além disso, elas também podem ser usadas para investigar os fluxos de conhecimento em uma comunidade epistêmica, como faz [Thagard \(1993\)](#) ao empregar métodos computacionais para investigar questões típicas da filosofia da ciência, como o processo de descoberta científica e a formação de conceitos teóricos.

A partir desta visão do papel operacional de conhecimento, podemos retomar a definição de [van der Grient \(2011\)](#) apresentada anteriormente. Tendo em vista o plausível fecho epistêmico das simulações, é difícil afirmar que uma simulação produza conhecimento que não esteja contido no conjunto original de consequências lógicas de um modelo, mas ela expande o subconjunto destas consequências que está ao alcance da comunidade epistêmica, efetivamente aumentando o conjunto das informações acionáveis pelos cientistas. Sob esta concepção, é possível falar que simulações computacionais produzem conhecimento científico, ainda que a justificação deste conhecimento seja contingente aos processos de verificação e validação ([BEISBART, 2017](#)).

2.4 Simulação e representação

Para que as simulações computacionais possam gerar informações acionáveis sobre o sistema-alvo, suas saídas devem representar de maneira adequada os estados observados desse sistema. Caso contrário, estamos diante da situação identificada por [Beisbart e Norton \(2012\)](#), em que a análise dos resultados da simulação traz informações não sobre o objeto que se pretende estudar, mas sim sobre o objeto simulacional em si. Para que as afirmações feitas a partir de uma simulação tenham qualquer pretensão de ser aplicáveis aos fenômenos estudados, precisamos entender o que constitui uma representação adequada do sistema-alvo.

Muitas das abordagens teóricas de simulações partem de alguma espécie de *realismo científico*, sustentando que o objetivo das teorias científicas é fornecer uma descrição correta do mundo ([OKASHA, 2016](#), p. 59), de forma que as entidades observáveis e não-observáveis envolvidas nestas teorias possuiriam referentes no “mundo real”. A esta posição, podem ser contrapostas visões *antirealistas*, para as quais a aceitação de uma teoria científica não significa que ela é uma descrição verdadeira do fenômeno, mas sim que possui virtudes intelectuais que a tornam interessante, como a adequação empírica ([VAN FRAASSEN, 1980](#), p. 10).

Simulações podem, em princípio, ser úteis tanto para uma investigação que parta de princípios realistas — caso no qual o objetivo do sistema simulacional é fornecer uma descrição verdadeira do sistema-alvo — quanto para perspectivas antirrealistas, que delimitarão virtudes intelectuais a serem seguidas na construção e interpretação das simulações.²¹ Para os fins desta discussão, tentaremos indicar como a simulação pode se relacionar com o alvo de sua descrição sem assumir posições mais específicas em relação à natureza deste alvo ou das teorias científicas.

Uma primeira tentativa de entender a natureza da representação que as simulações fazem de seu objeto de estudos é dizer que seu modelo estabelece uma *analogia* entre as saídas do sistema computacional e o sistema representado. A analogia é uma figura de linguagem bem estabelecida, na qual se estabelece uma correspondência entre aspectos de um sistema *A* e aspectos de um sistema *B*, que possui aspectos *positivos*, isto é, as propriedades de *A* aplicáveis a *B*; *negativos*, indicando as propriedades de *A* que não se aplicam a *B*; e *neutras*, cuja aplicabilidade é desconhecida ([GELFERT, 2016](#), p. 8). Para uma simulação, isto significa dizer que as dinâmicas entre os aspectos positivos implementados no modelo e os aspectos negativos presentes na sua construção, o mapeamento entre a simulação e o sistema alvo permitirá que as propriedades neutras sejam atribuídas

²¹ O debate entre realismo e antirrealismo científicos foge ao escopo desta dissertação, mas pode ser visto nas práticas de modelagem de autores como [Gilbert e Troitzsch \(2005, p. 17\)](#) e [Elsenbroich \(2012, p. 2\)](#), refletindo não uma propriedade necessária das simulações, mas sim compromissos valorativos e metodológicos prévios dos pesquisadores.

a um ou outro conjunto da analogia.

A noção de simulação enquanto analogia nos fornece uma primeira aproximação das estruturas simulacionais, mas é insuficiente para fundar os usos que desejamos. Uma primeira lacuna está na ausência de critérios para que uma característica seja ou não incluída na analogia; é possível construir uma analogia entre a simulação e o sistema-alvo que deixe de lado propriedades sabidamente relevantes para explicar o comportamento que desejamos estudar, tornando necessária a adição de outros critérios para estabelecer a utilidade da analogia.

A redução das simulações a analogias enfrenta um problema adicional quando estamos tratando de ciências sociais: ainda que o modo de raciocínio analógico seja comum nas ciências naturais,²² ele não necessariamente se adequa às formas reconhecidas de investigação de problemas sociais. Em campos mais adeptos de métodos formais, como alguns campos da filosofia política ou o *mainstream* da ciência econômica,²³ a analogia entre campos conceituais é mais aceita como fundamentação de argumentos, mas outros atribuem um papel secundário às analogias na construção teórica; um exemplo é a prática historiográfica tradicional, em que a analogia não é vista como elemento de prova (ELTON, 1983, p. 96–97), atuando apenas como fonte de estímulo para a investigação. Nestes casos, dizer que uma situação é análoga a outra — por exemplo, que uma crise de sucessão em um partido que governa um país é análoga a uma crise de sucessão dinástica em uma monarquia — não é visto como um sinal de que os fenômenos são regidos pela mesma lei ou mesmo por regras similares, mas sim como um fator que pode motivar o acadêmico que se interessa por um dos temas a estudar o outro.

A noção de correspondência entre a simulação e o sistema-alvo pode ser refinada a partir da importação de ferramentas da lógica formal. Uma abordagem, inicialmente proposta por Newton C. A. da Costa e Steven French (GELFERT, 2016, p. 19), expressa os modelos em termos de n -uplas de conjuntos, formadas por objetos e suas propriedades, quantidades (e funções associadas) e relações. Sobre estes conjuntos, é definida uma *estrutura parcial* que estabelece relações parciais entre estas n -uplas: cada relação R_i pode ser aplicável para algumas n -uplas, inaplicável para outras n -uplas e ter aplicabilidade indeterminada para um terceiro conjunto. Gelfert (2016, p. 19) destaca a correspondência entre essas estruturas parciais e a visão tripartite de analogias discutida acima, e o uso de relações entre conjuntos impõe uma estrutura formal para a relação analógica entre os sistemas, que são complementadas pelas exigências que a reprodução adequada do domínio impõe à formulação destas relações.

²² Gelfert (2016, p. 7) traz dois exemplos na prática científica: a descoberta por Kekulé da estrutura molecular do benzeno, que o folclore científico associa ao sonho deste pesquisador com o *ouroboros*, e a teoria da luz de Huygens, proposta por analogia ao modelo de ondas mecânicas.

²³ McCloskey (1998, p. 37) mostra a ubiquidade da analogia na retórica da ciência econômica, mesmo quando os argumentos analógicos não são empregados de maneira explícita.

Ao tratarmos a representação simulacional como um conjunto de isomorfismos parciais, temos as ferramentas conceituais para tratar simulações que se baseiam em múltiplos modelos conceituais, já que a estrutura parcial pode conter vários isomorfismos que sejam aplicáveis apenas a regiões específicas do domínio modelado. Estas simulações baseadas em múltiplos modelos podem usar modelos que descrevem um mesmo fenômeno em vários níveis — variando o nível de detalhamento de acordo com a relevância de regiões específicas do sistema — ou *ensembles* de modelos, em que diversas simulações são executadas em paralelo e o resultado é produzido a partir de algum método de agregação.

Winsberg (2010, p. 74–83) traz um exemplo de modelo em múltiplos níveis, empregado no estudo de fraturas em uma placa metálica. Simula-se o comportamento global da placa a partir de um modelo baseado em malhas, mas a região mais próxima da placa é simulada em um nível molecular e as moléculas do ponto de fratura seriam tratadas com um modelo em nível quântico. A adoção de múltiplos modelos permite que as regiões críticas sejam cobertas com um detalhamento que teria custo computacional proibitivo se estendido ao corpo como um todo, mas os modelos podem gerar resultados incoerentes entre si, exigindo o emprego de técnicas de integração que serão melhor descritas no Capítulo 4 deste texto.

Podemos enxergar as simulações multiníveis como um conjunto de isomorfismos parciais: cada modelo estabelece uma relação diferente entre suas saídas e os dados observacionais, relação esta que é aplicada à região do domínio em que o modelo é válido. Como os modelos são aplicáveis em domínios distintos, não há sobreposição entre as entradas de cada isomorfismo e, portanto, é possível construir trivialmente uma estrutura parcial a partir da união entre os níveis da simulação.

Já os *ensembles* de modelos possuem uma complicação adicional, uma vez que o resultado final será produzido pela agregação dos isomorfismos parciais correspondentes a cada modelo. Este agregado final pode ser em si tratado a partir de estruturas parciais, mas o tratamento dos isomorfismos correspondentes a cada modelo individual pode ser mais complexo, uma vez que pode haver sobreposição entre os domínios de isomorfismos pertencentes a modelos diferentes e que mapeiam esses pontos comuns a imagens distintas. A modelagem formal destes *ensembles* pode ser interessante para o entendimento da validade local dos modelos e para a sua harmonização, mas não será objeto de discussão aprofundada nesta dissertação.

2.4.1 Simulações, teoria e observações

As simulações cumprem o papel epistêmico a elas atribuído quando fornecem informações acionáveis. Esta acionabilidade pode ser entendida de um ponto de vista prático: um agente epistêmico, ao guiar sua conduta com base em uma informação qualquer, espera que esta informação não o induza a agir de maneira contraproducente a seus pró-

prios objetivos. Por exemplo, um cientista que use como sustentação de seus argumentos um modelo que o leva a conclusões rejeitadas por sua comunidade epistêmica dificilmente obterá sucesso em persuadir seus pares,²⁴ e uma empresa que baseie sua estratégia de marketing em simulações que não refletem o comportamento de seu público-alvo provavelmente terá prejuízos no mercado.

Em geral, as simulações computacionais são utilizadas para produzir informações sobre sistemas com existência independente do programa simulador.²⁵ Por isso, os modelos lógico-conceituais descritos na seção anterior buscam estabelecer relações entre as saídas das simulações e observações feitas do sistema-alvo da simulação, de forma a gerar expectativas sustentáveis sobre estes sistemas. Como discutido na Seção 1.3, a qualidade das expectativas geradas a partir das simulações pode ser assegurada a partir de processos de validação, que são dependentes de dados sobre o sistema estudado.

A natureza dos dados necessários para a construção e validação de uma simulação está relacionada com o domínio em que ela é aplicada. No Capítulo 4 teceremos considerações específicas sobre o processo de coleta de dados e a determinação das variáveis relevantes para um dado problema, mas antes precisamos entender de maneira mais precisa a natureza da informação sobre o sistema-alvo que é empregada na construção de uma simulação.

Em uma primeira passagem, podemos falar que estes dados são obtidos a partir da *observação* do sistema-alvo, seja ela feita de maneira direta, seja ela feita por meio de instrumentos de medida.²⁶ O processo de coleta de dados é ele próprio condicionado pelas teorias aceitas pelos responsáveis pela coleta, bem como pelos desdobramentos tecnológicos destas. Em um nível operacional, a escolha das variáveis relevantes para a modelagem de um problema — e, portanto, do que se observará no sistema-alvo — é reflexo das variáveis que os modelos teóricos existentes propõem como relevantes (MCCLOSKEY, 1998, p. 39). A dependência da teoria também se dá em um nível mais essencial: as concepções teóricas e os compromissos que as acompanham também acabam por definir quais são os objetos relevantes de estudo em um campo, bem como os meios tidos como aceitáveis

²⁴ Kuhn (2012) traz uma discussão já clássica sobre a forte adesão dos cientistas a compromissos metodológicos, instrumentais, conceituais e teóricos (seção IV), que estimulam a resistência a mudanças conceituais até que o paradigma vigente entre em crise (seção VIII). Em nosso exemplo, este conservadorismo da prática científica torna provável a rejeição do modelo “divergente” mesmo que ele aponte uma falha no modelo aceito até o momento; esta falha será percebida como um contra-exemplo pontual, a menos que revele uma anomalia grande no modelo vigente ou surja em um cenário no qual este já é objeto de desconfiança pela comunidade.

²⁵ Discutiremos, no Capítulo 3, usos em que as simulações não possuem uma relação óbvia com referentes externos e que, portanto, demandam outros critérios para acionabilidade. Tendo em vista o interesse desta dissertação no estudo de fenômenos sociais reais, a discussão se concentrará na relação entre uma simulação e seus referentes externos.

²⁶ Humphreys (2004, p. 122) relativiza esta distinção, afirmando que é importante para ciência não o que pode ser observado em princípio, mas sim os objetos e fenômenos detectáveis na prática, incluindo aí o uso de meios tecnológicos.

para investigar estes objetos (KUHN, 2012, p. 42).

Sob esta perspectiva, não seria possível falar em uma coleta de dados brutos, já que mesmo estes insumos da simulação computacional são percebidos a partir de um prisma teórico que delimita os dados observacionais, e a própria ideia de que dados de um determinado tipo são passíveis de observação pode ser alterada pela disponibilidade dos meios de coleta.²⁷

No entanto, esta antecedência da teoria em relação à observação tem seus limites: de acordo com Brewer e Lambert (2001, S177–S180), os arcabouços teóricos existentes realmente influenciam a percepção dos dados observacionais, mas não ao ponto de determiná-los totalmente. Nesta perspectiva cognitiva, o sentido da observação seria construído de uma maneira *bottom-up* que pode ter ambiguidades, cuja interpretação é feita a partir da perspectiva *top-down* da teoria. O impacto da teoria não seria forte o suficiente, contudo, para sobrepujar o que Brewer e Lambert (2001, S184) caracterizam como fortes evidências observacionais. Como alertava Schönemann (1994, p. 158), não é possível desacoplar a teorização formal dos dados empíricos, já que a interpretação destes não é inteiramente condicionada pelas construções abstratas *a priori* que marcam essa formalização.

Uma consequência dessa subdeterminação dos dados observacionais pela teoria é a possibilidade de descoberta de novos fenômenos e de contra-exemplos à teoria.²⁸ As observações podem, em princípio, chocar-se com as informações de um sistema teórico, o que afastaria a hipótese de que usar estes dados observados para validar o modelo seja um raciocínio circular.²⁹ Porém, mesmo um impacto limitado da teoria na percepção já indica que a percepção do sistema-alvo é mediada por aspectos teóricos. Como a acionabilidade das informações geradas a partir de uma simulação depende da relação entre esta e o sistema-alvo, a qualidade do conhecimento produzido depende da qualidade da modelagem subjacente à simulação e das teorias que guiam a coleta de dados.

²⁷ Arthur Jensen (1980, p. 172) traz o exemplo de como a temperatura, antes da invenção do termômetro, era vista como um exemplo de grandeza imensurável.

²⁸ Sob uma perspectiva kuhniana, anomalias são reconhecidas na atividade científica diante de uma falha sistemática e persistente por parte das ferramentas de um paradigma em resolver os quebra-cabeças que este delimita no espaço científico. Assim, as simulações em si, cuja validade é também avaliada a partir de critérios internos ao paradigma, não trariam em si a identificação de contra-exemplos, mas, na resolução destes quebra-cabeças, podem apontar o pesquisador na direção de problemas difíceis de serem superados e que, eventualmente, podem apontar para a existência de uma anomalia na visão de mundo vigente (KUHN, 2012, seção VI).

²⁹ Brewer e Lambert (2001) apontam para o impacto da teoria em processos de coleta de dados, interpretação e análise dos resultados, que seria maior que o verificado na observação em si. Na Seção 4.4 discutiremos algumas propostas de ordem prática para tentar mitigar a circularidade de validar teorias a partir de dados por elas moldados.

2.5 O status científico das simulações

Simulações computacionais bem-desenhadas podem atuar como uma fonte de justificação de uma crença, na medida em que o seu projeto e os seus resultados sejam compatíveis com fontes de conhecimento já estabelecidas. A justificação oriunda das simulações tem suas raízes nos processos que as constroem, funcionando como uma generalização que não pode justificar proposições que não sejam consequência lógica do modelo usado na construção do programa. Ainda assim, a simulação pode contribuir para a investigação de seu objeto, seja de forma indicial, ao sugerir pontos para os quais deve ser direcionada a pesquisa científica, seja de forma operacional, ao expandir o espaço de informação acionável sobre um fenômeno.

É possível, portanto, extrair conhecimento sobre um fenômeno social a partir da análise dos resultados das simulações computacionais. Esta validade deriva do modelo teórico empregado no desenvolvimento, bem como da relação entre a implementação e o sistema-alvo, que pode ser representada a partir de estruturas parciais entre elementos representados na linguagem da teoria dos conjuntos. A combinação entre estas estruturas aumenta a confiança na analogia entre a simulação e o sistema que ela pretende representar, justificando as inferências feitas a partir daquela.

3 Formalização de fenômenos sociais

O olhar percorre as ruas como se fossem páginas escritas: a cidade diz tudo o que você deve pensar, faz você repetir o discurso, e, enquanto você acredita estar visitando Tamara, não faz nada além de registrar os nomes com os quais ela define a si própria e todas as suas partes.

Italo Calvino, *As Cidades Invisíveis*: As Cidades e os Símbolos 1

Apresentamos no Capítulo 1 algumas aplicações acadêmicas e não-acadêmicas de simulações computacionais a domínios sociais. Tendo em vista a diversidade de métodos, objetos e recortes teóricos empregados por acadêmicos e não-acadêmicos, não adotaremos aqui uma definição fechada da noção de *fenômeno social*, de forma a não excluir do escopo de aplicação das simulações computacionais objetos de estudo tidos como válidos por comunidades científicas ou técnicas específicas. Assim, consideraremos como fenômeno social qualquer fenômeno envolvendo interações de indivíduos, grupos ou sociedades.

Nosso objetivo aqui é investigar as possibilidades de aplicação de simulações em campos influenciados pelas ciências sociais. Adotar uma abordagem prescritiva para descrever a noção de “fenômeno social” poderia deixar de fora fenômenos que as comunidades científicas ou técnicas interessadas consideram como válidas, além de ser um desafio que foge ao escopo desta dissertação. Analisaremos, pois, as formas como as simulações são de fato empregadas na pesquisa das ciências sociais e nas aplicações práticas por ela influenciadas.

A construção de uma simulação de um dado fenômeno social, como discutido nos capítulos anteriores, envolve a representação do objeto simulado por de um modelo computacional. Este modelo expressa em uma linguagem formal os atores, interações e eventos tidos como relevantes para representar de maneira adequada o fenômeno estudado, o que é usado como base para justificar inferências realizadas a partir desse modelo formal. Um exemplo clássico de modelo formal é a descrição de um fenômeno a partir de equações quantitativas,¹ mas esta formalização também pode ocorrer em modelos qualitativos, a partir do uso das ferramentas da lógica e da teoria de conjuntos (GOERTZ; MAHONEY, 2012, p. 11).

O debate sobre o uso de métodos quantitativos e qualitativos nas pesquisas em ciências sociais possui longa história,² representada em uma gama de posições que

¹ Por exemplo, as equações diferenciais de Lanchester, utilizadas na modelagem de forças relativas de exércitos em conflito (DAVIS, P. K., 1995, apêndice).

² Este debate pode ser visto como parte de um debate mais amplo sobre a existência de diferenças

vão desde a rejeição da pertinência da quantificação³ à visão de que abordagens qualitativas seriam uma forma particular das explicações quantitativas,⁴ passando por toda sorte de posições intermediárias. As simulações computacionais não forçam, em princípio, a preferência por abordagens quantitativas ou qualitativas: pode-se usar o programa simulacional para gerar vários cenários de saída, cujos resultados são analisáveis por técnicas estatísticas, ou para a construção de alguns poucos cenários que sejam estudados a partir das técnicas qualitativas para estudos de casos. De fato, é possível empregar simulações como uma combinação entre as abordagens qualitativas e quantitativas, ao fazer uso de *insights* extraídos a partir de ambas para a construção dos modelos de simulação.

Venturini, Pablo Jensen e Latour (2015, p. 1.3–1.6) discutem o uso das simulações como uma ferramenta para investigar o vínculo entre macromodelos de comportamento social, para os quais se pode coletar informações agregadas, e micromodelos de interações de indivíduos. Os resultados produzidos nestes dois níveis de modelagem são comparados ao longo de múltiplas execuções de um programa simulacional, e esta possibilidade de geração de vários cenários simulados gera dados em quantidade maior do que seria possível através da observação de fenômenos relativamente raros ou ainda não verificados.

A simulação permitiria, então, que os modelos de comportamento de agentes individuais fossem inspecionados em busca de contradições internas e externas;⁵ estas, surgidas da comparação dos resultados dos modelos locais com o comportamento global esperado. Por exemplo, um modelo baseado em agentes que busque explicar um fenômeno econômico será avaliado tanto a partir das interações entre os agentes individuais — o *micromodelo* do sistema — quanto em termos dos padrões econômicos que podem ser observados a partir da população como um todo — o *macromodelo*.

Embora este papel integrador das simulações tenha surgido em um contexto

entre o método das ciências naturais, solidamente quantitativas, e as ciências humanas; para Berlin (1974), a ideia de que estes dois campos de estudo seriam essencialmente diferentes encontra raízes em Giambattista Vico, pensador do século XVIII, culminando nos hermenêutas do início do século XX.

³ Um exemplo pode ser visto em von Hayek (1975), para quem as ciências sociais seriam capazes apenas de uma predição de padrões distinta da predição precisa fornecida pelas ciências naturais.

⁴ Para King, Keohane e Verba (1994, p. 230), os métodos quantitativos e os métodos qualitativos são baseados na mesma lógica de inferência, e as diferenças entre as abordagens seriam de natureza estilística. Já Goertz e Mahoney (2012, p. 220) falam na existência de dois paradigmas distintos de pesquisa social — um paradigma quantitativo, fundado em abordagens estatísticas, e um paradigma qualitativo, cuja formalização residiria na lógica e na teoria de conjuntos — que podem coexistir tanto dentro de uma mesma ciência quanto em um mesmo programa de pesquisa, a partir da adoção dos chamados métodos mistos (GOERTZ; MAHONEY, 2012, p. 226).

⁵ A preocupação habitual, aqui, é com a verificação de contradições lógicas entre as afirmações factuais extraídas dos resultados de uma simulação e as expectativas geradas pelos axiomas de teorias aceitas pela comunidade epistêmica. Uma alternativa é a formulação que Elster (1978, cap. 4–5) faz das *contradições reais*, em uma formulação em lógica formal da ideia de contradição presente na dialética de Hegel, que por sua vez serviu de base para as formulações marxianas (ELSTER, 1978, p. 3). Estas contradições não seriam meramente uma expressão de relações lógicas entre proposições, mas a tradução de dinâmicas individuais ou sociais que se referem a objetos com sua existência própria.

de relativa escassez de dados,⁶ Conte et al. (2012, p. 337–338) propõem que o desenvolvimento de técnicas de *big data* permite que estas simulações sejam validadas a partir de conjuntos volumosos de dados, gerando previsões micro e macro que podem ser testadas com experimentos ou com a análise dos dados gerados *in vivo* pelo sistema que se deseja estudar. A simulação desempenharia, então, o papel de auxiliar no entendimento do sistema ao direcionar o emprego de outros métodos.

O uso de modelos formais também pode colaborar com o diálogo entre tradições acadêmicas ao forçar a explicitação de alguns dos termos implícitos envolvidos nas construções teóricas, como discutido na Subseção 2.4.1. A tradução dos conteúdos teóricos para uma linguagem formalizada permitiria a expressão das noções teóricas em uma formulação passível de análise por pesquisadores externos a uma dada tradição científica. No momento, há baixa adesão às simulações dentre os estudiosos de fenômenos sociais, que impede que elas possam ser consideradas como uma *lingua franca* (LENHARD; KÜPPERS; SHINN, 2007) científica. Mesmo assim, seu uso para a expressão formal de teorias já sinaliza um meio potencial para *tradução* entre linguagens científicas associadas a distintos referenciais teóricos, uma vez que o processo de implementação e uso permite a identificação de termos e expressões em que há choque de interpretações ou que não pertencem ao jargão de uma das culturas (KUHN, 2012, p. 200).

Modelos formais estão no cerne de uma simulação computacional, não só por especificar o funcionamento do programa simulador como por influenciar a interpretação teórica deste modelo, seja por parte dos usuários da simulação, seja por parte de pesquisadores alheios ao arcabouço teórico empregado em sua construção. Tais modelos direcionam a pesquisa baseada em simulações,⁷ mas não necessariamente forçam a adoção de métodos alheios a uma dada tradição de pesquisa. Em vez disso, o uso de métodos formais pode ser feito como uma extensão dos métodos já reconhecidos de pesquisa social,⁸ como sustentaremos a seguir.

⁶ O que motiva o argumento de Venturini, Pablo Jensen e Latour (2015), discutido na Seção 1.4, de que as simulações perderiam relevância com a quantidade maciça de dados gerada pela moderna sociedade informacional. Esta abundância de dados exige de fato uma mudança no uso das simulações, priorizando algumas das aplicações que discutiremos nesse capítulo, mas não exaure o papel destas ferramentas computacionais, que passam a ser instrumentos de validação de hipóteses em vez de uma fonte de dados que agora podem, em muitos casos, ser obtidos por observação.

⁷ Elton (1983, p. 119, tradução própria) alerta para o perigo de validação autorreferencial das simulações: “Modelos ditam os termos de referência, definem os parâmetros, direcionam a pesquisa, e portanto são muito propensos a perverter a busca por evidência empírica ao torná-la seletiva.” No Capítulo 4 discutiremos técnicas que podem mitigar ou evitar essa distorção da evidência empírica.

⁸ Neste sentido, ver Elster (1978, cap. 2) e sua formalização de análises sociais construídas a partir de influências hegelianas e marxianas.

3.1 Simulações e os modos de explicação social

As simulações computacionais de fenômenos sociais são um tipo de modelo formal que pode contribuir para a organização das abordagens teóricas usadas para estudar um fenômeno social, seja esse estudo feito de forma qualitativa, ao capturar de maneira detalhada o comportamento dos objetos observados e suas regras lógicas, seja ele baseado em uma abordagem quantitativa, fornecendo dados que serão usados para análises estatísticas do comportamento simulado. Em ambos os casos, o uso de técnicas de simulação pretende contribuir para a *explicação* dos fenômenos observados em um plano social.

Uma explicação de um fenômeno social é a busca por motivos envolvidos na ocorrência ou não de um determinado fenômeno social (VON WRIGHT, 2004, p. 134). Para von Wright (2004, p. 132–134), a atividade de explicação não se confunde com o entendimento de dados comportamentais, mas é dependente desta para a reinterpretação dos fatos e até mesmo para a delimitação do objeto a ser explicado.

Para parte da literatura, explicações seriam a identificação de mecanismos causais (ELSENbroich, 2012, p. 7), enquanto outros autores veem a explicação como “*uma escolha entre estratégias interpretativas*” (WHITE, 2008, p. 14, ênfase nossa), dentre outras possibilidades. Os critérios de explicação podem ser distintos entre campos de estudos ou mesmo entre correntes metodológicas em uma mesma área,⁹ refletindo abordagens que diferem entre si não só por questões como a maior ou menor adoção de métodos quantitativos, mas principalmente em questões de valores de pesquisa, na definição de quais são as perguntas válidas para a investigação dos fenômenos e dos critérios que as respostas-candidatas devem atender para serem aceitas como válidas.¹⁰ O uso de simulações como parte do ferramental investigativo em ciências sociais será adequado na medida em que as técnicas simulacionais puderem ser adequadas aos modos de explicação aceitos pelas comunidades de pesquisa em ciências sociais.¹¹

Em sua análise sobre a noção de explicação nas ciências sociais, Mahajan

⁹ Um exemplo pode ser visto no debate entre Fogel (1982) e Elton (1983), que envolve em parte a pertinência das questões levantadas por historiadores quantitativos para a produção historiográfica, dado que ambos percebem estas questões como distintas daquelas levantadas por historiadores de abordagem “tradicional”, como o próprio Elton.

¹⁰ Esta construção faz uso da *matriz disciplinar* proposta por Kuhn (2012, p. 180–186) como refinamento da noção de paradigma científico. Ainda que se parta da ideia de que as ciências sociais são metodologicamente distintas das ciências naturais, a análise kuhniana ainda pode ser útil para identificar fatores que afetam as práticas dos membros dessas comunidades epistêmicas.

¹¹ Uma outra possibilidade seria afirmar que as simulações são peças de um novo modo de explicação de fenômenos sociais, como faz Fredkin (2003) na física ao propor a substituição de modelos matemáticos tradicionais por uma *mecânica digital* que trata os processos físicos de uma maneira informacional e digitalizada. Sem afastar, em princípio, esta possibilidade, concentraremos nossos esforços na conciliação de abordagens simulacionais com abordagens estabelecidas, subscrevendo à tese de Frigg e Reiss (2009) de que simulações não introduzem *per se* problemas filosóficos novos. A abordagem que adotamos também tem a vantagem, de um ponto de vista prático, de não dificultar o diálogo com áreas em que as simulações ainda não são parte do ferramental habitual do pesquisador.

(2011) identifica quatro modos explicativos aceitos por diferentes tradições acadêmicas. O primeiro deles é uma abordagem *causalista*, que trata os fenômenos observados como consequência de eventos anteriores necessários e suficientes para a ocorrência observada (MAHAJAN, 2011, p. 1). Os demais modos explicativos identificados por Mahajan (2011) deixam de lado a busca por causas mecânicas de eventos, julgando outras construções como mais adequadas para a explicitação dos motivos que desembocam em um determinado evento; podemos, então, chamá-los de *explicações não-causais*.

O segundo modo de explicação, a *explicação de razões para a ação*, busca identificar os motivos que levaram um agente a atuar de determinada forma (MAHAJAN, 2011, p. 27); de forma relacionada, porém distinta, o modo *hermenêutico* percebe as condutas de um agente a partir da expressão de sua visão do mundo, que atribui significação à ação (MAHAJAN, 2011, p. 96). Por fim, o modo *narrativo* tenta identificar a situação e as ações que deram margem à ocorrência do evento em questão (MAHAJAN, 2011, p. 95).

A busca por explicações de fenômenos sociais é um objetivo associado às investigações científicas, mas que também é relevante para os usos práticos de simulações. Por meio da representação computacional simulada de agentes sociais, uma empresa pode fundamentar suas expectativas em relação à conduta de outros atores de mercado, ou um governo pode obter subsídios para o desenho de políticas públicas (GRÜNE-YANOFF; WEIRICH, 2010, p. 44). Os resultados destas simulações “práticas” serão entendidas com o auxílio dos modelos teóricos empregados em sua construção, e podem fornecer explicações tanto para comportamentos passados inesperados quanto para justificar uma tomada de decisão voltada ao futuro. Em ambos os casos, a simulação será utilizada para a produção de explicações a partir dos modos verificados na pesquisa acadêmica, de acordo com as necessidades de persuasão dos *stakeholders* a quem a simulação se dirige.

3.1.1 Simulações como ferramenta de explicação causal

Explicações causais de fenômenos sociais tentam fornecer uma sequência de passos relevantes que culminam num estado observado (ELSENBROICH, 2012, p. 2). Uma explicação causal completa de um estado S_n a partir de um estado S_1 precisaria incluir todos os eventos que transformam S_1 nos estados subsequentes S_2, \dots, S_n , e estabelecer estas transformações como condições necessárias e suficientes para a situação final. Para King, Keohane e Verba (1994, p. 81–82, tradução própria), pode-se entender essas condições a partir da noção de *efeito causal*, que seria:

(...) a diferença entre a componente sistemática das observações feitas quando a variável explicativa assume um valor e a componente sistemática das observações comparáveis quando a variável assume outro valor.

Na prática, não é possível observar simultaneamente todos os valores que uma variável explicativa poderia tomar.¹² Além da dificuldade de observação dos resultados, também há o problema de identificação dos fatores relevantes, tendo em vista a complexidade dos fenômenos estudados. No Direito Penal, por exemplo, a *proibição de regresso* surgiu como um limitante da pesquisa causal (ZAFFARONI et al., 2010, p. 175), ao tornar condutas inofensivas irrelevantes para estabelecer a imputação criminal: alguém que meramente abre a porta de uma loja para um assaltante não será responsabilizado, em geral, pelo crime cometido, ainda que a entrada do agente na loja seja condição necessária para a ação delituosa.¹³

A dificuldade em identificar os fatores causais relevantes para uma explicação pode levar ao abandono ou ao refino desta noção. Variantes do primeiro caminho podem ser identificadas na proposta de Kelsen (1967, p. 76) de fundar a sua *Teoria Pura do Direito* em um princípio de imputação distinto da causalidade que este autor associa às ciências naturais e, um tanto antes, no esvaziamento da causalidade histórica promovido por Herder (WHITE, 2008, p. 85), dentre outros tratamentos metodológicos. Estas perspectivas divergentes precisam, pois, fornecer leituras alternativas da noção de explicação, ou propor um objetivo diferente para seus campos de estudo.

Explicações causais podem ser tornadas operacionais a partir de um relaxamento das condições de causalidade. Um caminho para isso, visível em King, Keohane e Verba (1994, p. 76), é enxergar a causalidade como uma construção teórica que será observada na prática como uma variável aleatória. Sob esta perspectiva, simulações podem contribuir para a medida dos efeitos causais ao proporcionar múltiplas amostras para o cálculo dos parâmetros relevantes, o que, partindo de um desenho adequado da simulação como descrito no Capítulo 4, melhora a estimativa quantitativa do efeito causal.

Outra abordagem é a busca por *mecanismos* causais, isto é, conexões entre eventos que não constituem uma história causal completa, mas apontam para a atuação dos fatores causais propostos no resultado sob estudo (ELSENbroich, 2012, p. 2). Neste caso, alguns dos estados s_i que seriam necessários para uma explicação causal completa não estão presentes, mas o mecanismo proposto fornece uma abordagem plausível para que não seja necessário preencher explicitamente estas lacunas. Uma simulação pode contribuir para a descoberta destes mecanismos ao fornecer um ambiente em que candidatos a fatores causais podem ser testados através de sua inclusão nas regras de simulação. Se

¹² Usando o exemplo de King, Keohane e Verba (1994, p. 76) um distrito eleitoral em um sistema bipartidário pode eleger um ou outro partido, e o efeito associado ao caminho não-tomado não será observado. Discutiremos mais adiante no capítulo a relação entre esta observação e o uso *contrafactual* das simulações.

¹³ A teoria causalista do delito, em suas múltiplas variantes, buscava identificar o ato que causava uma conduta típica, efetuando a partir daí as operações lógicas para verificação da existência ou não de um crime. Esta perspectiva foi adotada no Código Penal de 1940 e persistiu até a reforma da Parte Geral deste código realizada em 1984, quando foi substituída por uma perspectiva *finalista*, que tinha como uma das principais promessas clarificar tal imputação (ZAFFARONI et al., 2010, p. 164 ss.).

os resultados obtidos com a simulação se mostrarem incompatíveis — seja de um ponto de vista qualitativo, seja do ponto de vista quantitativo — com as observações do fenômeno ou com o arcabouço teórico em que se trabalha, a pretensão causal do mecanismo proposto é esvaziada.

3.1.2 Simulações e as explicações não-causais

A aplicação do modo de explicação causal aos fenômenos sociais encontra, além das dificuldades de ordem prática na identificação de relações causais em cenários complexos, objeções de natureza conceitual. Já no século XVIII, Giambattista Vico defendia como essencial ao entendimento histórico a compreensão da visão de mundo dos agentes humanos, que não poderia ser reduzida ao entendimento científico dos fatos e documentos relevantes (BERLIN, 1974, p. 132–133). Esta visão, posteriormente reformulada por uma tradição com raízes no já mencionado Herder (BERLIN, 1974, p. 139), apontaria para uma clivagem essencial entre os métodos das ciências naturais e os métodos das ciências sociais, que inviabilizaria a investigação puramente causal nos domínios destas. Uma segunda linha crítica pode ser identificada a partir de von Wright (2004, p. 136–138), para quem a busca por causalidade em ciências sociais, ainda que não seja impossível, serviria na verdade identificar vínculos entre estados ou eventos do mundo, cujas explicações demandariam recursos além da conexão causal.

Qual seria, então, o modo de explicação pertinente para o estudo dos fenômenos sociais? Uma resposta possível é o deslocamento do objeto: em vez de explicar eventos com existência independente, as ciências sociais investigariam as *ocorrências de ação humana no mundo social* (MAHAJAN, 2011, p. 27).¹⁴ A ação seria um comportamento humano movido pela intenção de atingir um determinado resultado, e explicar as consequências desse comportamento no mundo social seria identificar as razões que motivaram o agente.¹⁵

O diagnóstico das razões para a ação pode envolver o uso de técnicas de simulação computacional, que podem ser usadas para modelar o comportamento dos agentes relevantes e avaliar os fatores envolvidos na tomada de decisão em cenários simulados. De particular valia para este uso são as simulações baseadas em agentes normativos, que são um tipo particular de agente artificial cuja conduta é vinculada tanto por normas internas — por exemplo, regras fixadas por sua programação — quanto por normas externas. A interação entre estes dois níveis de regras faz com que o agente precise ponderar várias

¹⁴ Zaffaroni *et al* 2010, 88 destacam o *modelo teleológico* da “ação racional conforme a fins” empregada por Weber na explicação social, bem como o papel que a noção de ação desempenha no sistema jurídico-penal (ZAFFARONI *et al.*, 2010, p. 100).

¹⁵ Na sistemática do Direito Penal, é visível o abandono de um sistema causal de ação — que considerava a ação como a conjunção entre um movimento corpóreo e um resultado unidos pela relação de causa e efeito (ZAFFARONI *et al.*, 2010, p. 84) — em favor de um sistema finalista que enfatiza não o processo causal, mas a vontade com finalidade que motiva a ação potencialmente delituosa (ZAFFARONI *et al.*, 2010, p. 89). Há, pois, uma pretensão de classificar a conduta como delito a partir das razões que a guiam.

regras, possivelmente conflitantes entre si, para decidir suas ações, de forma que a construção de simulações de agentes normativos pode ser útil tanto para a investigação dos agentes individuais quanto do impacto social de suas condutas (ELSENbroich; Gilbert, 2013, p. 199–201).

Explicações baseadas no *entendimento hermenêutico* também se preocupam com o significado dado pelo agente a uma conduta, mas situam essa interpretação no contexto histórico dos agentes (MAHAJAN, 2011, p. 50). A produção de sentido de uma ação, sob esta perspectiva, não é meramente fruto de decisão individual, mas surge em um ambiente cultural que define as práticas e valores tidos como relevantes para aquela comunidade; a explicação das ações dependeria não só do conhecimento de fatos e eventos, mas da empatia com a visão de mundo que as move (BERLIN, 1974, p. 137). O processo investigativo forneceria um *entendimento* do significado da ação em questão, sem se preocupar em explicar suas consequências (MAHAJAN, 2011, p. 75).

Qualquer explicação nesta perspectiva depende de uma reconstituição complexa de contexto, que pode ser difícil ou até mesmo impossível de representar inteiramente de forma computacional. Isto não significa, no entanto, que simulações nada tenham a ver com uma investigação no modo hermenêutico. Por um lado, as preocupações com o contexto das construções teóricas podem ser empregadas para a construção de modelos mais sofisticados, através da incorporação dos elementos viáveis de implementar nas condições do problema. Por outro lado, simulações de modelos de comportamento dos agentes relevantes podem permitir a visualização mais clara dos processos interpretativos envolvidos no entendimento do contexto social e histórico; de especial valia para este fim são os modelos que buscam replicar atividades cognitivas, que podem fornecer uma forma de inspeção de processos mentais.

Já as explicações *narrativas* de fenômenos sociais tratam o evento a ser explicado como uma consequência das ações de indivíduos e grupos em uma determinada situação histórica (MAHAJAN, 2011, p. 81). A explicação narrativa apresenta um evento como determinado por um complexo de condições e ações, que é reconstruído de maneira racional pela atividade de pesquisa (MAHAJAN, 2011, p. 87). Uma narrativa, portanto, atribui sentido a um evento particular a partir do comportamento dos agentes inseridos em um contexto durante um intervalo de tempo específico, tratando dos fatores específicos daquela ocorrência.

Simulações computacionais podem ser harmonizadas com a busca por narrativas explicativas. O objetivo deste emprego não é a execução de múltiplos cenários para a computação de variáveis descritivas, mas a descrição das dinâmicas em uma única execução da simulação. Por exemplo, as simulações podem dar margem a uma análise detalhada de um cenário contrafactual definido a partir da configuração de seus parâmetros iniciais de execução; uma simulação dinâmica permite o acompanhamento da transformação do

cenário ao longo do tempo de execução, o que serve de base para avaliar as reações dos agentes à evolução de seu contexto.

O uso de simulações computacionais de fenômenos sociais, por sua natureza formal, possui forte associação com as propostas causais de explicação. Isso não significa, no entanto, que não existam margens para a aplicação destas técnicas na construção de outras formas explicativas, em especial ao viabilizar novas abordagens ou a investigação de novos problemas.¹⁶ Em vez de rejeitar *prima facie* o emprego de simulações como uma forma de colonização metodológica, pesquisadores adeptos de modos não-causais de explicação podem, assim, usá-las para aperfeiçoar suas análises de formas condizentes com os pressupostos que já adotam.¹⁷

3.1.3 Explicação e Predição

Simulações podem ser empregadas também com fins *preditivos*: dado um modelo que se mostrou confiável em seu treinamento, seus parâmetros podem ser modificados para tentar gerar resultados futuros de um fenômeno. Por exemplo, um modelo que descreva adequadamente o comportamento de consumidores de um setor pode ser usado como um preditor da reação desses clientes a uma mudança no modelo de negócios da empresa. As inferências feitas a partir dessa projeção são, por sua vez, validadas pela qualidade das explicações fornecidas pelo modelo.

Em uma perspectiva clássica do processo científico, conseguir prever a ocorrência de um evento exigiria o conhecimento dos fatores que o geram, ou seja, explicá-lo (ELSENbroich, 2012, p. 2). Haveria uma *simetria* entre as estruturas lógicas que explicam um fenômeno e as estruturas que permitem a predição de suas ocorrências futuras, estruturas essas que difeririam apenas em sua orientação temporal: enquanto a explicação trata de eventos passados, a previsão é orientada para o futuro (ELSENbroich, 2012, p. 2).

A conexão entre explicação e predição é relativizada nas explicações não-causais: conhecer as razões que levam um agente a agir de uma determinada forma no instante t_1 pode não ser suficiente para explicar a conduta t_2 , ainda mais se a própria dimensão temporal é relevante para a construção de sentido da ação, e raciocínios análogos se aplicam ao modo hermenêutico e ao modo narrativo. Mesmo dentro de uma visão causal dos fenômenos sociais, é possível construir simulações que expliquem os mecanismos de

¹⁶ Elton (1983, p. 116, tradução própria) destaca o papel de técnicas computacionais na expansão do espaço de problemas válidos para a investigação histórica: “[a história ‘científica’] cumpre melhor sua função verdadeira quando ela soma às técnicas disponíveis para a análise de evidência histórica, principalmente para que problemas que historiadores de outrora não podiam ou não iriam estudar possam ser sujeitos a investigações históricas.”

¹⁷ Fogel (1982) destaca uma propensão ao sincretismo metodológico na pesquisa histórica: “Historiadores cada vez mais tomaram de cada uma das [ciências] sociais as ideias que poderiam adicionar potência e profundidade às suas análises, sem se comprometer com uma visão totalizante do comportamento humano ou da evolução histórica”. A adoção de técnicas de simulação pode proceder do mesmo modo, expandindo o ferramental conceitual sem forçar um cientificismo metodológico.

funcionamento de um sistema-alvo sem que elas possam fornecer base para predições de comportamentos futuros, como nos modelos baseados em agentes ([ELSENBROICH, 2012](#), p. 6), ou obter preditores que não forneçam explicações interpretáveis, por exemplo, ao incorporar técnicas de redes neurais ([YARKONI; WESTFALL, 2017](#)).

Distinguir entre o papel explicativo e o papel preditivo de simulações não significa, necessariamente, optar por um ou outro, já que em muitos casos práticos as duas abordagens podem ser conciliadas. A ênfase por uma ou outra será determinada pelos usos ao qual se destina a simulação, bem como pelas prioridades da comunidade de usuários.

3.2 Papéis analíticos das simulações

Simulações computacionais são, em geral, construídas com propósitos instrumentais: as explicações, as predições, ou outros usos se encaixam na resolução de um problema acadêmico ou prático. A partir dos diversos casos de uso trazidos nos capítulos anteriores, pode-se esboçar uma taxonomia das simulações de fenômenos sociais, o que de fato foi feito por trabalhos como o manual de [Gilbert e Troitzsch \(2005, p. 5–6\)](#) e o *survey* de [Grüne-Yanoff e Weirich \(2010\)](#). Partindo dessas propostas classificatórias, discutiremos alguns dos papéis que simulações podem desempenhar na análise de fenômenos sociais.

Um uso comum de simulações é a *aproximação*: em cenários em que é impossível ou dispendioso obter uma solução fechada de um modelo representado por um sistema de equações, pode-se obter uma solução aproximada através da agregação de resultados de várias instâncias de uma simulação. [Humphreys \(2004, p. 150\)](#) traz o exemplo dos modelos Ising no ferromagnetismo, cuja formulação analítica envolve uma integral que, mesmo em domínios simples, pode ter 10^6 graus de liberdade. Para lidar com esta alta dimensionalidade, soluções práticas costumam fazer uso de algoritmos de Metropolis, baseadas nos métodos de Monte Carlo citados na Subseção 2.1.2, para fornecer soluções aproximadas a partir de amostragem estatística.

Para [Grüne-Yanoff e Weirich \(2010, p. 30\)](#), o exemplo de [Humphreys \(2004\)](#) não seria uma simulação do fenômeno, mas sim um método de cálculo estocástico sem pretensão de representação do sistema-alvo do simulador. Esta diferença de classificação é, no fundo, fruto de uma diferença nas definições de simulação: o algoritmo de Metropolis em si não é uma representação do fenômeno simulado, mas é um elemento constitutivo de um sistema computacional que se encaixa na concepção adotada a partir da Seção 1.2, já que usa a execução temporal do algoritmo probabilístico para resolver de maneira aproximada as equações que descrevem computacionalmente o fenômeno de estudo.

Ao excluir os modelos de Monte Carlo de sua concepção mais estrita de simulações, [Grüne-Yanoff e Weirich \(2010, p. 30\)](#) estabelecem um contraste destes com o que

chamam de *propósito mimético* de uma simulação. Nessa perspectiva, as simulações em sentido estrito seriam construídas com a intenção de imitar características do sistema-alvo, reproduzindo processos relevantes para a análise em tela em um processo que combina a *imitação* de algo concreto e a *criação* de cenários imaginários, cuja composição dependerá de técnicas que mediem estes dois aspectos (GRANSCHKE, 2017). Tal preocupação com os processos internos do fenômeno simulado serviria como fonte de validade tanto das explicações quanto das previsões baseadas em uma simulação mimética, cujo comportamento seria análogo ao de seu objeto de estudo.

A imitação fornecida por uma simulação mimética pode encontrar alguns limites. Uma primeira fonte de problemas reside na modelagem: se os modelos e as condições iniciais da simulação não forem minimamente plausíveis para a comunidade epistêmica em questão, a informação produzida a partir dela será de utilidade questionável frente aos membros dessa comunidade. Em particular, a discussão sobre aceitação dos modelos em geral passa pelo grau de *idealização*¹⁸ envolvida em uma simulação: o modelo conceitual dos objetos de uma simulação necessariamente envolverá a abstração de algumas das características salientes do objeto representado; se os aspectos abstraídos forem particularmente relevantes para a descrição que um determinado referencial teórico faz do objeto em questão, a simulação dificilmente será compatível com esta teoria.

Ainda que as idealizações envolvidas no modelo sejam tidas como válidas por seus destinatários, é possível que o cenário simulado esteja fora da região de aplicabilidade, o que esvaziaria as pretensões de validade da simulação: uma simulação do orçamento federal brasileiro construída dentro das regras da Constituição de 1988 produzirá resultados anômalos se for usada para analisar o gasto público sob a Constituição de 1946, ainda que calibrada a partir dos valores adequados para as variáveis macroeconômicas relevantes, já que os dois textos constitucionais definiam diferentes obrigações estatais. Por fim, detalhes de construção da simulação podem prejudicar a imitação do sistema-alvo, o que torna necessários os cuidados de construção que discutiremos no Capítulo 4.

Um subtipo dos modelos miméticos é o *contrafactual*, em que se calibra o modelo para imitar um sistema com base em seu comportamento histórico, mas se introduz um *ponto de divergência* em algum ponto do passado. A partir deste ponto, as condições são alteradas em relação às verificadas nas observações do sistema-alvo, o que forneceria uma explicação indireta da relevância das variáveis alteradas ao apontar como as mudanças alteram ou não os resultados esperados. Para Elster (1978, p. 176), as afirmações contrafactuais estariam intimamente ligadas à busca por causalidade — seja através de uma definição contrafactual de causalidade, seja a partir da implicação de cenários contrafactuais a partir de explicações causais —, mas a análise do que não aconteceu pode também servir para fortalecer explicações não-causais ao fornecer contraexemplos a pro-

¹⁸ Esta noção será discutida em mais detalhes na Subseção 4.1.4.

postas alternativas de explicação, viabilizando o que Fogel (1982) descreve como a *análise de sensibilidade* dos argumentos postos.

O exercício contrafactual é próximo de duas outras práticas simulacionais: a *ficção* e a *previsão*. Nos três casos, os resultados gerados pela simulação não podem ser diretamente associados a algum estado observado do sistema-alvo no momento da execução; assim, sua validação é dependente não de uma correspondência ao mundo externo, mas sim da coerência de suas regras de funcionamento com os referenciais teóricos empregados e da relevância das divergências introduzidas na modelagem. O uso ficcional das simulações, que será abordado nesta seção, distingue-se dos outros dois por não possuir um referencial observável sequer em princípio, enquanto os resultados de uma previsão podem ser validados externamente com a passagem do tempo. Já o contrafactual não tem esta facilidade, mas sua extrapolação de cenários não-ocorridos se baseia nas regras verificadas a partir das observações do sistema-alvo antes do ponto de divergência.

A principal diferença entre o contrafactual e a previsão reside na possibilidade de observação dos resultados previstos, já que não é possível repetir na prática as condições contrafactuais criadas. Estes dois modos, todavia, compartilham algumas preocupações em sua construção, como a necessidade de validar os resultados obtidos através do uso de parte dos dados observados como conjunto de referência e a preocupação com a distância temporal: conforme o modelo contrafactual se afasta do ponto de divergência — ou o preditor, do instante presente —, os efeitos das diferenças entre o modelo e o sistema representado são potencializados, prejudicando a representação do sistema. Este efeito pode ser identificado — e por vezes até atenuado — através de múltiplas execuções da simulação, que podem ser comparadas entre si para a validação dos resultados.

Existem situações nas quais o construtor da simulação não é plenamente familiar com o fenômeno a ser modelado ou com as consequências de seus pressupostos teóricos, o que pode prejudicar a qualidade da imitação de uma simulação mimética. Ainda assim, as simulações podem contribuir para o entendimento de um sistema a partir da *exploração*¹⁹ dos modelos científicos envolvidos em sua construção, através da geração de casos que apontem os limites e insuficiências de um modelo ou da variação de elementos do modelo conceitual, como as premissas de modelagem, os elementos representados ou as relações entre estes elementos conceituais.

Gelfert (2016, p. 83–94) apresenta uma descrição de quatro possíveis funções para os modelos exploratórios, que são aplicáveis às técnicas de simulação, já que estas possuem em seu cerne um modelo lógico-conceitual do fenômeno simulado. Um primeiro uso é como *ponto de partida* para a construção de modelos mais sofisticados: as insuficiências

¹⁹ Destaca Gelfert (2016, p. 74–74) que a noção de exploração pode ser entendida em dois sentidos: uma exploração *específica*, que busca explicações para um dado estímulo e converge em torno desse fato, e uma exploração *diversificante*, em que a resposta cognitiva é orientada pela busca de novidades. Ambas as formas são acessíveis através de simulações, como discutiremos.

de uma simulação inicial apontarão pontos que precisam ser refinados nas iterações seguintes de modelagem. Este uso fornece uma base para o segundo uso possível: simulações podem implementar *explicações candidatas* para um dado fenômeno, que serão avaliadas pela comparação entre elas e também com o modelo de base. Simulações também podem servir como *provas de conceito* da aplicabilidade de uma abordagem metodológica ou de um processo específico a um dado problema.²⁰ Por fim, as simulações podem ajudar a fazer o *recorte do objeto* estudado, identificando as condições iniciais e de contorno relevantes. Nestes processos, as simulações contribuem não só para o refinamento das hipóteses e técnicas de modelagem, mas também abrem possibilidades de investigação (ELTON, 1983, p. 120) relativas aos próprios arcabouços teóricos usados em sua construção.

Simulações podem também desempenhar um papel didático, ao evidenciar o impacto de regras do sistema simulado e dos valores específicos das variáveis. Este aspecto didático pode ser subsidiário aos usos descritos anteriormente, conforme os resultados da simulação são empregados para melhorar o entendimento sobre o sistema, mas pode também ser o objetivo único da simulação. A simulação das dinâmicas de segregação social feita por Schelling (1972) pode ser percebida por esse prisma: além de desempenhar o papel de *prova de conceito* de simulações, ao mostrar que um modelo simples pode dar origem a padrões espacialmente segregados, ela também foi criada com o intuito de mostrar *como* um computador poderia ser utilizado na realização do modelo de segregação que o autor anteriormente desenvolveu.

Os usos discutidos ao longo desta seção estão relacionados primariamente com o papel analítico das simulações, ou seja, o uso delas como ferramentas de estudo de cenários postos, partindo de premissas e conjuntos de dados fechados. As simulações, contudo, não são objetos isolados do mundo que pretendem estudar, pois seus resultados podem influenciar o comportamento de entidades relevantes para o cenário estudado.

3.3 Papéis retóricos das simulações

Simulações são empregadas como argumento para a sustentação de discursos sobre fenômenos sociais, tanto em um contexto científico quanto em contextos extra-acadêmicos. Seja no uso mimético, seja no uso preditivo, ou em qualquer outro modo de uso, os resultados de uma simulação trazem informações que invocam para si alguma validade científica a partir da qualidade da representação que fazem do fenômeno-alvo. Essa pretensão de validade, por sua vez, é usada para sustentar discursos científicos, como a validade da teoria simulada, ou práticos, como uma decisão corporativa de investimento

²⁰ Gelfert (2016, p. 86) traz o exemplo do modelo ecológico de Lotka-Volterra, que não só mostrou que equações diferenciais seriam aplicáveis à dinâmica de populações como também propôs um mecanismo para explicar como oscilações das populações de presas e predadores podem surgir mesmo sem influências externas.

motivada por uma projeção simulada de resultados futuros.²¹ Em todos estes casos, a simulação é usada para persuadir os seus destinatários a alterar suas posições com base na interação entre sua visão de mundo e a *refiguração*²² dos resultados simulacionais (GRANSCHKE, 2017, p. 37–40).

O uso das simulações como elemento de um discurso científico fica explícito no modo narrativo de explicação que discutimos anteriormente, no qual os resultados da simulação são utilizados como elementos para a construção de uma história plausível que explique os fenômenos estudados. A criação desta narrativa se vale explicitamente da confiança no modelo computacional e na sua implementação para justificar porque esta narrativa, e não alguma das alternativas, descreve melhor o cenário em questão. Mesmo em outros modos de explicação, o emprego de simulações é feito como forma de legitimar os argumentos teóricos feitos a partir de seus resultados, permitindo que a comunidade epistêmica avalie se tais afirmações são razoáveis ou não dentro do contexto teórico-metodológico em que a simulação é usada (MCCLOSKEY, 1998, p. 150). A simulação se torna, assim, parte do estilo argumentativo de um dado texto, como faz Robert Fogel ao simular cenários alternativos para sustentar sua tese de que as ferrovias não ocuparam papel central no desenvolvimento econômico dos EUA (MCCLOSKEY, 1998, p. 80).²³

Além das questões de adequação ao método discutidas na seção sobre modos de explicação, um fator crucial para entender o uso argumentativo das simulações é a inteligibilidade da simulação. Um sistema computacional que permita a análise de várias dimensões dos resultados, bem como dos parâmetros e das relações envolvidas na execução do programa, está aberto ao escrutínio da comunidade científica, que pode realizar testes e alterações no modelo para verificar sua compatibilidade com o estado da arte de uma determinada corrente metodológica no campo de estudo em questão.²⁴ A possibilidade de entender o modelo implementado serviria, então, para aumentar a confiança nos resultados produzidos.

O papel da inteligibilidade das simulações na validação dos discursos é bem

²¹ Um exemplo é a máquina de Phillips, construída pelo economista neozelandês Bill Phillips como um modelo hidráulico da economia britânica pós-Segunda Guerra Mundial, usando a variação no nível de água com corante para mostrar o impacto de políticas econômicas na demanda agregada (RYDER, 2009).

²² O conceito de refiguração explorado por Gransche (2017) é extraído da obra de Paul Ricoeur, na qual é utilizado para designar a transformação da experiência temporal do leitor de um texto a partir da interação com uma narrativa (GAGNEBIN, 1997).

²³ Em seu livro *Railroads and American Economic Growth*, publicado em 1964, Fogel faz um contraste entre a proposição de que o estabelecimento das ferrovias resultou no aumento do rendimento nacional e o impacto hipotético de substitutos potenciais — como canais para transporte hídrico —, fazendo uso de fontes históricas para obter estimativas dos custos envolvidos nestes cenários alternativos. A partir deste contrafactual, Fogel conclui que o ganho anual de rendimento devido às ferrovias foi de 0,6% até 1890, valor que não condiz com a ideia de um efeito concreto significativo (MCCLOSKEY, 1998, cap. 5).

²⁴ A implementação prática deste ponto depende do acesso aos dados, ao código da simulação e às condições em que ela foi executada, condições que nem sempre se verificam na prática.

mais ambíguo que o retrato apresentado acima. Se, por um lado, o entendimento de um modelo pode contribuir para aceitação do uso de seus resultados, por outro pode expor os limites e inconsistências deste modelo. Críticas à modelagem realizada ou aos parâmetros empregados enfraquecem a pretensão de validade de uma simulação, ao questionar a qualidade de sua representação do fenômeno a ser descrito. Os dados utilizados para a construção e a validação de uma simulação podem ser atacados como inadequados ou mesmo insuficientes para reproduzir os resultados que o usuário original afirma ter obtido. E, mesmo que este processo de replicação consiga produzir resultados compatíveis com aqueles obtidos pelos usuários originais da simulação, os dados gerados permitirão a realização de outras análises, cujos resultados podem divergir do uso original das simulações. A presença de uma ou mais dessas situações pode ser o suficiente para desfazer os ganhos retóricos que motivaram a aplicação original da simulação.

O pleno entendimento de uma simulação pode enfraquecer ou fortalecer sua credibilidade, a depender da aceitação das hipóteses e teorias envolvidas, mas a dificuldade em destrinchá-lo pode acabar por afastar críticas ao seu emprego. Um modelo pode ser aceito mesmo sem ser inteligível, caso faça uso de abordagens metodológicas já aceitas pela comunidade a que os resultados simulacionais se destinam. Nestes casos, um observador que analise a simulação pode acabar por aceitá-la mesmo que sem fazer um exame adequado do modelo subjacente, ou até mesmo descartar críticas pertinentes a aspectos da simulação como mero entendimento incorreto de um modelo opaco. Um exemplo concreto pode ser visto no caso *K. W. vs. Armstrong* em Idaho, nos Estados Unidos, em que a justiça federal americana entendeu que órgãos públicos estavam tomando decisões com base em um modelo matemático que produzia resultados errôneos, mas que não era questionado por uma combinação de inércia burocrática e presunção de correteude das fórmulas complexas construídas por consultores externos ([STANLEY, 2017](#)).

A falta de estímulo para questionar modelos que, no caso geral, são aceitos pela comunidade epistêmica em questão pode ser motivada tanto por aspectos sociais — como a tentativa de não expor ignorância sobre aspectos tidos pela comunidade como fundamentais — quanto por uma falta de reflexão metodológica a respeito da pertinência dessa instância específica do método para o problema em questão. Em ambos os casos, a própria dificuldade em entender a simulação acaba servindo como argumento para que as inferências nela baseadas sejam aceitas, em uma prova por intimidação metodológica. O engajamento crítico com as simulações computacionais exige, pois, que não só os construtores das simulações como também os seus consumidores sejam educados a entender as ideias por trás das simulações computacionais e os aspectos performáticos envolvidos nos processos miméticos e na apresentação de resultados ([GRANSCHKE, 2017](#)).

Simulações computacionais de fenômenos sociais desempenham um papel argumentativo que vai além do mero fornecimento de dados para argumentos científicos

ou que fazem uso de teorias das ciências sociais. Os elementos formais que constituem a simulação serão peças de argumentos, e sua construção e divulgação podem ser elementos usados contra ou a favor da posição que a simulação foi usada para sustentar, ou mesmo como peça de argumentos pouco relacionados ao uso original.

3.4 Os usos das simulações

O uso de técnicas computacionais, incluindo as simulações, em domínios típicos das ciências sociais pode encontrar resistência de parte da comunidade científica relevante. Muitas vezes, tal resistência é fundamentada na visão de que estas técnicas seriam relevantes para a identificação de relações causais, o que seria uma abordagem das ciências naturais que não é adequada a domínios sociais. Ainda que existam correntes metodológicas em diversos campos — em especial a economia — que são menos refratárias à investigação causal, o uso de simulações é compatível com outros modos de explicação, como a construção de narrativas ou a busca por razões que guiam a conduta de atores relevantes. Assim, a simulação seria uma ferramenta compatível com as abordagens teórico-metodológicas dominantes nas ciências sociais.

Uma das principais utilidades da simulação na investigação de fenômenos sociais é a possibilidade de avaliar teorias sem precisar aguardar a ocorrência de um cenário específico ou realizar um experimento deliberado. Isto não quer dizer, no entanto, que as simulações não produzam efeitos nos sistemas que pretendem investigar: as conclusões extraídas a partir da análise de seus resultados podem ser empregadas como fonte de legitimação de argumentos não só em um contexto científico mas também em discursos políticos, econômicos ou de outras esferas sociais — por exemplo, uma simulação que mostre que o aumento de pena de um crime reduziu as taxas de criminalidade em um dado cenário pode ser utilizada como justificativa para mudanças na política criminal.

As simulações não fornecem ao estudioso de fenômenos sociais um método que realize as pretensões de objetividade tipicamente associadas com o uso de métodos formais nas ciências sociais. São, no entanto, compatíveis com os arcabouços conceituais destes campos de estudo, e seus múltiplos usos podem contribuir para lançar luz sobre aspectos de fenômenos sociais que não são adequadamente explicados pelas abordagens tradicionais, enriquecendo o estudo destas questões e, por extensão, as possibilidades extra-acadêmicas fundadas nestes estudos.

4 Desenvolvimento de simulações

Marco Polo descreve uma ponte, pedra por pedra.

— *Mas qual é a pedra que sustenta a ponte?, pergunta Kublai Khan.*

— *A ponte não é sustentada por esta ou aquela pedra, responde Marco, mas pela curva do arco que estas formam.*

Kublai Khan permanece em silêncio, refletindo. Depois acrescenta:

— *Por que falar das pedras? Só o arco me interessa. Polo responde:*

— *Sem pedras o arco não existe.*

Italo Calvino, *As Cidades Invisíveis*: As cidades e os mortos 1

Os capítulos anteriores discutiram as condições em abstrato para que simulações possam representar fenômenos sociais. Todavia, a mera possibilidade de construção de uma representação não é suficiente para a maioria das aplicações que descrevemos até o momento: é preciso que a simulação proposta no caso de uso seja adequada para os usos desejados. Mesmo uma solução teoricamente adequada pode ser prejudicada por problemas nos dados que a fundamentam ou na sua execução em um sistema computacional concreto, e as próximas páginas discutirão as formas como estes problemas de ordem prática podem ser identificados e evitados ou mitigados.

A construção de uma simulação computacional pode ser entendida como um *processo* mais ou menos estruturado. Gilbert e Troitzsch (2005, p. 18–26) descrevem o processo de construção de uma simulação científica como dividido em quatro etapas: o *desenho* do modelo do sistema-alvo, a *construção* do modelo desenhado, a *verificação e validação* das saídas e a *publicação* dos resultados obtidos a partir do modelo. Esta estrutura mais simples captura aspectos essenciais da atividade de pesquisa a partir de simulação, mas deixa de lado elementos que podem ser relevantes para outros usos das simulações.¹ Além disso, as simulações utilizadas para fins não-acadêmicos muitas vezes são construídas para uso contínuo — e não apenas durante a fase de coleta de dados de um artigo, como pode ocorrer na pesquisa acadêmica —, o que demanda atividades para a manutenção do programa simulacional construído.

Já Siegfried (2014, p. 13–14) traz um ciclo de modelagem de sete etapas: depois de uma *fase preliminar* voltada à identificação das necessidades do projeto, há a *definição*

¹ Em particular, a publicação da simulação e de seus resultados em usos extra-acadêmicos pode envolver questões de propriedade intelectual, segredos comerciais e difusão institucional de conhecimento que não serão abordadas nesta dissertação.

do problema de maneira estruturada, que serve de base para a *análise do sistema real*. Esta etapa tem como produto um modelo conceitual do sistema, que é então *formalizado e implementado*, produzindo uma simulação executável, que será configurada a partir de um processo de *experimentação*. Por fim, o resultado da simulação configurada serve para a *interpretação* a partir de conhecimento especializado. Cada uma das etapas pode revelar questões que serão resolvidas a partir de alterações em passos anteriores, e pode se valer de procedimentos já estabelecidos de engenharia de software, como descreve o autor.

O ciclo de desenvolvimento de simulações proposto por Siegfried (2014) traz uma estrutura operacional que é interessante para a produção sistemática de simulações, mas toma como objetivo último a produção de simulações com objetivos bem-definidos de representação mimética do sistema-alvo. Na prática, nem sempre as simulações são usadas nestas condições, já que é possível construir simulações sem que haja uma noção clara de finalidade do sistema, e o objetivo simulacional pode ser não a reprodução de um objeto-alvo, mas a formação de um modelo conceitual, em um uso exploratório que não é só uma etapa intermediária de construção.²

Tendo em vista as limitações de ambos os fluxos discutidos, propomos aqui uma divisão que conserve as características essenciais deles. A partir de uma primeira etapa de *coleta de dados* que sirvam como base sobre o fenômeno que se deseja estudar, é feita a *modelagem formal* do sistema-alvo, com base no ferramental teórico disponível ao construtor. Este modelo é então *implementado* como um programa de computador, que passará por um processo de *calibração* utilizando os dados coletados e só então será *executado* para gerar os resultados desejados, cuja *análise* fornecerá informações acionáveis sobre o sistema-alvo. No caso de simulações que sejam executadas em vários momentos ou de maneira contínua, há também a necessidade de *manutenção* do programa, para a eliminação de eventuais defeitos identificados e a atualização para o tratamento de dados concretos.

Este modelo, contudo, simplifica bastante o processo ao impor uma sequência linear de fases. Na prática, os estágios podem ser executados em ordem distinta da prescrita — como discutido na Subseção 2.4.1, a coleta de dados não se dá em um cenário pré-teórico, mas já traz em si pressupostos embutidos, e mais dados podem ser coletados ao longo do desenvolvimento do sistema —, bem como afetar a execução de estágios anteriores e posteriores. A análise do fluxo, portanto, deve levar em conta que a comunicação entre os estágios não é unidirecional, investigando as múltiplas possibilidades de interação entre cada estágio do ciclo de construção das simulações, bem como o refinamento da simulação ao longo de várias iterações deste ciclo.³

² Sobre o tema, ver nossa discussão na Seção 3.4.

³ De Marchi (2005, p. 137–140) traz a *autodestruição* como um dos principais méritos das simulações: a construção de uma simulação levaria à identificação de seus problemas e à construção de modelos melhores, em um processo de refinamento e reforço dos modelos existentes sobre um fenômeno.

Muitos dos problemas de ordem prática identificados a partir deste fluxo de desenvolvimento de simulações encontram análogos nos processos de *engenharia de software*, que fornecerão técnicas aplicáveis ao desenvolvimento. Outra fonte de técnicas úteis é a *ciência de dados*, cujas preocupações com a coleta, o tratamento e a análise de dados são muitas vezes aplicáveis no tratamento das simulações. Na prática, pode ser complicado identificar como aplicar estas ferramentas já existentes na construção das simulações, e o objetivo deste capítulo é fornecer substratos para que elas sejam empregadas de maneira produtora e compatível com os objetivos identificados para as simulações computacionais.

4.1 Modelagem formal de sistemas

Modelos formais de fenômenos sociais ocupam um papel central na construção de simulações computacionais destes objetos de estudo. Para representar de maneira adequada o seu sistema-alvo, estas simulações podem fazer uso de técnicas como modelos de Monte Carlo,⁴ modelos baseados em agentes, modelos de teoria de filas e modelos de teoria dos jogos, dentre outras técnicas.⁵ Cada uma destas abordagens possui suas peculiaridades, que são objetos de literatura específica.⁶ Nesta seção, o foco será discutir questões gerais de modelagem que possam balizar a escolha dos modelos para um dado problema, ainda que a decisão final seja determinada pelo entendimento teórico que se tem das características do domínio de estudo ou pelas competências e recursos disponíveis para os construtores.

Um modelo científico de um fenômeno é um objeto físico ou conceitual que representa as propriedades deste, através dos processos descritos no Capítulo 3 desta dissertação. Tal representação compara o modelo a um *sistema-alvo*, que é uma construção conceitual que abstrai algumas características do fenômeno estudado e o descreve em termos que viabilizam a comparação direta com o modelo (WEISBERG, 2013, p. 90–95). A prática de modelagem fornece, assim, um mapa do fenômeno, que pode ser usado tanto para a predição e identificação de características-chave desse objeto de estudo quanto para a descoberta de novos fenômenos relevantes (MILLER; PAGE, 2007, p. 36–40).

A construção de modelos para simulações pode se valer de práticas já estabelecidas para a modelagem científica. Para De Marchi (2005, p. 121), uma primeira etapa

⁴ Definidos no Capítulo 1 e objeto da discussão da Subseção 2.1.2.

⁵ No Capítulo 1 apresentamos uma seleção, sem pretensões de exaustividade, de técnicas de simulação já empregadas no estudo de fenômenos sociais. Além destas, nada impede, em princípio, a adaptação de outras abordagens ou o desenvolvimento de novas técnicas relevantes.

⁶ O manual de Gilbert e Troitzsch (2005) traz um panorama das técnicas de simulação empregáveis em domínios sociais, como modelos microanalíticos, simulações multiníveis, autômatos celulares, modelos multiagentes e modelos de aprendizado de máquina, trazendo os fundamentos destas abordagens e indicações de materiais para aprofundamento, ainda que não aborde avanços recentes como *deep learning* e a difusão da computação distribuída, posteriores a seu lançamento.

na produção de modelos computacionais é a *codificação* do problema, isto é, a definição de uma representação para os elementos relevantes; o autor traz o exemplo de uma simulação baseada em teoria dos jogos, na qual seria relevante codificar os agentes, as estratégias e as funções de utilidade. A partir de tal codificação, o modelo precisaria definir uma *busca por soluções* na população, utilizando técnicas de otimização estocástica para buscar boas estratégias para o problema modelado. Ainda que nem todos os usos de simulações computacionais possam ser reduzidos a problemas de otimização,⁷ esta divisão entre a expressão do problema em uma linguagem adequada para a simulação e a definição dos processos computacionais a serem executados nos oferece um ponto de partida, de onde podemos buscar as expressões computacionais mais adequadas ao tipo de problema em questão e aos objetivos da construção da simulação.

Uma proposta de fluxo de modelagem com aplicação mais geral é apresentada por Cioffi-Revilla (2014, p. 232–233). Aqui, o primeiro passo seria diagnosticar as *questões de pesquisa* que o modelo pretende resolver, sejam elas puramente acadêmicas ou de natureza mais aplicada. A formulação destas questões estaria sujeita a constritores de ordem prática, como a falta de dados,⁸ a incompletude do referencial teórico, ou limitações de prazos, recursos ou pessoal, que podem limitar o alcance da pesquisa. Além destes pontos levantados por Cioffi-Revilla (2014), uma simulação pode ser alterada depois de sua construção para investigar novas questões, ou mesmo ser empregada em usos que não foram previstos originalmente, o que altera os objetivos de pesquisa;⁹ ainda assim, estas adições posteriores estarão dependentes do modelo construído a partir das questões originais, o que torna o diagnóstico inicial especialmente relevante para a modelagem.

O segundo estágio do processo definido por Cioffi-Revilla (2014, p. 233–234) seria o *desenho conceitual* do modelo, através da escolha de aspectos a abstrair do fenô-

⁷ Por exemplo, a formação de narrativas ou o tratamento de agentes que se pautam por regras não-utilitárias podem ser difíceis de expressar como problemas clássicos de otimização.

⁸ Em resposta a este problema é que a coleta de dados foi posicionada por nós como primeira etapa do ciclo de construção de uma simulação, ainda que ela muitas vezes seja um trabalho transversal que se inicia antes da modelagem e continua mesmo depois que a simulação está operacional.

⁹ O estudo de fatores não-técnicos pode se beneficiar de abordagens de engenharia de software como as presentes em Sommerville (2011). Além dos aspectos comuns ao desenvolvimento tradicional de software, existem também fatores específicos deste domínio, como as *culturas de modelagem*, formadas por profissionais que usam modelos a partir de práticas e pressupostos metodológicos compartilhados. A diferença entre culturas pode fazer com que simulações não produzam os objetivos desejados por seus construtores ou por seus usuários, na medida em que elementos dos modelos ou os modelos em si são interpretados de formas distintas e que podem levar a resultados inesperados. O estudo das culturas de modelagem não é abordado nesta dissertação, mas é objeto de literatura especializada, da qual destacamos como porta de entrada os artigos em que MacKenzie e Spears (2014b,a) analisam o uso dos modelos de cópula gaussiana no mercado financeiro. Esta família de modelos foi adotada pelos bancos de investimento para o tratamento de derivativos, mas era um elemento “estranho” à cultura de modelagem dessa área (MACKENZIE; SPEARS, 2014a), e a extensão de modelos de cópula a estes novos domínios, de maneira metodologicamente heterogênea, contribuiu para a crise de 2008, na medida em que os êxitos iniciais das aplicações do modelo levaram as instituições financeiras a deixar de lado problemas na adoção destes modelos que pareciam secundários em um primeiro momento (MACKENZIE; SPEARS, 2014b).

meno. O processo de abstração, aponta Weisberg (2013, p. 57), determinará quais são as propriedades *focais* de um modelo, isto é, aquelas cuja reprodução acurada é essencial para que o modelo atinja seus objetivos. Em uma simulação computacional, estas propriedades serão representadas de maneira *explícita* ou *implícita* por variáveis que armazenam os valores relevantes à sua descrição, que pode ser feita de maneira *discreta* ou *contínua*.

Uma variável particularmente relevante para as simulações computacionais de fenômenos sociais é o tempo. Em muitos dos usos de simulação descritos nos capítulos anteriores, o programa simulacional não obtém uma resposta estática, mas sim uma descrição de um processo que ocorre ao longo de um intervalo de tempo, como a simulação feita por Schelling (1972) da segregação em uma população. Como estes modelos podem gerar em minutos ou segundos a reprodução de fenômenos que ocorrem ao longo de anos ou gerações, como é o caso da segregação espacial, torna-se necessária a distinção entre o *tempo de execução* do programa simulacional e o *tempo interno* à lógica do modelo.

O tempo interno do modelo pode ser tratado de maneira implícita: cada segundo de execução — o dito “tempo de parede” — corresponde a um intervalo de tempo, não necessariamente constante, no interior do modelo. Outra alternativa é que o tempo seja tratado como uma variável explícita do programa, cujo valor varia de maneira *contínua* — por exemplo, em uma simulação de eventos em tempo real — ou *discreta*, como seria o caso em um jogo de xadrez, no qual cada jogada ocorre em um turno específico. A escolha do tratamento do tempo no modelo terá consequências para a implementação do modelo computacional, mas pode também ter impactos no modelo conceitual, ao influenciar, por exemplo, o tratamento de eventos que ocorrem ao mesmo tempo na lógica interna ao modelo.

A escolha das abstrações deveria, em princípio, ser feita a partir das questões que se deseja responder, mas na prática acabará influenciada pelos recursos computacionais e pelas técnicas disponíveis aos construtores da simulação (CIOFFI-REVILLA, 2014, p. 234), bem como os constritores não-técnicos que afetam a etapa anterior. Para a identificação destes constritores, os desenvolvedores de simulações podem se beneficiar não só do conhecimento próprio, mas também fazer uso de consultas a especialistas externos e até mesmo envolver os *stakeholders* externos que serão afetados de maneira direta ou indireta pela simulação. Ao envolver estes atores, os projetistas de uma simulação conseguem não só antecipar problemas éticos e jurídicos que possam surgir da simulação como também se beneficiar do conhecimento que estas pessoas ou organizações possuem sobre o sistema que se pretende simular.¹⁰ Uma vez que as decisões feitas neste estágio condicionarão as etapas de implementação do sistema computacional, a incorporação de diversos pontos de vista pode ser particularmente relevante para evitar alterações que serão custosas se

¹⁰ Para ambas as finalidades, o uso de abordagens como o *design participativo* (DAVIS, J., 2009) pode ajudar projetistas de sistemas computacionais a incorporar uma diversidade perspectivas ao projeto da simulação, como discutido por Almada e Attux (2018).

realizadas em etapas posteriores do ciclo de desenvolvimento.

4.1.1 Modelos probabilísticos

A escolha do modelo adequado para uma simulação depende das características do fenômeno específico que se pretende estudar. Em alguns casos, como o estudo do impacto que um prêmio de loteria tem na vida do beneficiado, o fenômeno estudado já contém em si uma fonte de aleatoriedade. Em outros, como uma simulação baseada em métodos de Monte Carlo, o ferramental da probabilidade é introduzido como uma forma de aproximar resultados determinísticos.

Em um modelo puramente determinístico de um fenômeno, o estado da simulação é uma função do estado inicial e da sequência de operações aplicadas sobre este estado. Por exemplo, um jogo de xadrez pode ser reconstituído a partir das posições iniciais das peças e dos movimentos registrados em notação algébrica. A partir de um mesmo estado inicial, aplicar as mesmas operações necessariamente produzirá os mesmos resultados, a menos de erros na execução do processo; assim, executar novamente uma simulação puramente determinística só é interessante como técnica de validação ou para verificar o impacto da variação de parâmetros do modelo.

Uma forma de tratar a aleatoriedade em modelos de simulação é considerar que as probabilidades relevantes independem do tempo, como é o caso em um sorteio de números para uma loteria, que segue sempre as mesmas regras. Existem casos, no entanto, em que esta independência não se aplica: se quisermos saber a probabilidade de um indivíduo receber sozinho o prêmio máximo dessa loteria, precisaremos ver não só a probabilidade de uma dada sequência de números ser sorteada, mas também fatores como o número de pessoas que participam de cada sorteio da loteria e eventuais padrões na distribuição dos números jogados. Estes fatores podem variar com o tempo — por exemplo, sorteios de fim de ano ou com prêmios acumulados atraem maior demanda de jogadores não-habituais —, o que exige cuidados na modelagem.

O tratamento matemático das probabilidades variantes no tempo é feito através de *processos estocásticos*. Um processo estocástico pode ser definido como uma família de variáveis aleatórias X_t , em que t é uma variável de índice que pode ser discreta ou contínua. Processos estocásticos são caracterizados por seu espaço de estados — que contém os valores possíveis das X_t , por seu conjunto de índices T e pelas relações de dependência entre as variáveis X_t (PINSKY; KARLIN, 2010, p. 4). Uma forma de processo estocástico é o *processo de Markov*, em que, conhecido o estado atual do sistema, ter informação sobre estados prévios não altera o conhecimento sobre as probabilidades dos estados futuros: por exemplo, o tamanho de uma fila pode ser representado a partir do tamanho atual dessa fila, o número de pessoas que chega em um dado instante e o número de pessoas que saem dessa fila (PINSKY; KARLIN, 2010, p. 79–92).

A cada instante da execução de uma simulação baseada em modelos estocásticos, o programa amostra as variáveis aleatórias relevantes para computar o novo estado do modelo. Como esta amostragem pode produzir qualquer dos valores possíveis das variáveis aleatórias, duas execuções de uma mesma simulação podem produzir resultados distintos. Se a execução A da simulação ocupou no instante t um estado $S_{t,A}$ distinto do valor $S_{t,B}$ verificado em uma execução B posterior no mesmo instante, é possível que os estados para todos os instantes posteriores a t sejam diferentes. Um exemplo seria um modelo de sucessão dinástica, no qual o reconhecimento ou não do direito de um herdeiro ao trono pode afetar a linha sucessória por gerações. Neste caso, execuções singulares podem ser relevantes para análise qualitativa, feita através da análise dos *logs* de estados intermediários e do estado final da simulação, mas podem ter resultados que não são representativos do comportamento esperado do sistema. Para obter uma análise mais robusta e generalizável, pode ser necessário executar o modelo várias vezes com o mesmo parâmetro, gerando assim vários “universos possíveis” que serão analisados em conjunto (GILBERT; TROITZSCH, 2005, p. 25) para ver quais resultados se aplicam a todos os cenários ou a frações relevantes destes.

4.1.2 Objetivos de modelagem

Simulações são construídas com objetivos específicos de pesquisa, ainda que estes possam ser modificados com o passar do tempo. Para responder questões de investigação específicas, a modelagem enfatizará aspectos diferentes do fenômeno que se pretende estudar, deixando de lado outras características que possam ser menos relevantes. Cioffi-Revilla (2014, p. 249–305) faz uma distinção entre duas modalidades de simulações: os *modelos orientados a variáveis* — nos quais a dinâmica temporal do sistema-alvo é capturada a partir de equações formais, como nos modelos baseados em teoria de filas — e os *modelos orientados a objetos* — nos quais a simulação se baseia nas menores unidades relevantes e descreve o comportamento geral de forma emergente a partir delas, como é o caso nas simulações baseadas em agentes.

Outra distinção relevante é em relação ao *alvo* da modelagem. Uma modelagem com alvo específico representa um sistema específico e se propõe a responder questões sobre este sistema. A partir da construção ou adoção de uma estrutura lógica do sistema e dos critérios de interpretação para a representação do sistema-alvo, a simulação pode servir de base para uma *análise completa* que elucide propriedades globais do modelo e de seus estados ou para a investigação de hipóteses mais restrita, respondendo questões específicas sobre o sistema-alvo a partir do comportamento da simulação (WEISBERG, 2013, p. 74–97).

A modelagem que não é direcionada a um alvo específico (WEISBERG, 2013, p. 114–134) pode adotar como alvo um fenômeno geral — Weisberg (2013, p. 115) traz

o exemplo de descrições das origens da reprodução sexuada —, implementar uma hipótese explicativa de fenômeno a partir de alvos sabidamente não-existent ou impossíveis (WEISBERG, 2013, p. 121–129), ou mesmo trabalhar com um alvo em aberto a partir das especificações iniciais (WEISBERG, 2013, p. 129–131). Em todos estes casos, as simulações desempenham um papel exploratório do espaço de investigação, partindo de pressupostos bem-definidos mas sem possuir uma linha de chegada clara que forneça critérios de aceitação ou rejeição para a simulação e para a análise de seus resultados.

4.1.3 Complexidade e reducionismo

Uma das principais justificativas práticas para o uso de simulações computacionais na investigação de fenômenos sociais é que as técnicas computacionais ofereceriam ferramentas adequadas para lidar com a *complexidade* destes fenômenos. Um sistema complexo pode ser entendido como um sistema composto por um grande número de componentes¹¹ que interagem entre si (SIEGFRIED, 2014, p. 11), como é o caso em muitos sistemas que modelam o comportamento de indivíduos.

Segundo Cioffi-Revilla (2014, p. 211–212), podemos identificar a complexidade em fenômenos sociais a partir de três características:

1. A *racionalidade limitada* dos agentes envolvidos, que não agem de forma perfeitamente racional por obstáculos de computação, por vieses cognitivos e pela informação imperfeita;
2. A *quase-decomposibilidade* destes sistemas em subsistemas e módulos quase autôntidos que interagem entre si; e
3. A *emergência* de ordem no sistema a partir de decisões locais e descentralizadas.

A emergência é uma característica que recebeu particular atenção nas discussões sobre a complexidade. Uma perspectiva para seu tratamento é o *reducionismo*, segundo o qual o comportamento de sistemas poderia ser explicado a partir do comportamento de suas partes constituintes. Em última instância, padrões sociais poderiam ser reduzidos a problemas de física de partículas, ainda que a agregação destas partículas leve a sistemas com comportamento distinto, que exigiria uma função de transição entre estes níveis de modelagem (MILLER; PAGE, 2007, p. 41–42).

Contrapondo-se a estas posições, a ideia de comportamento social emergente afirma que certos fenômenos só podem ser descritos com a adição de novas categorias:

¹¹ A definição adotada por Siegfried (2014) considera que os componentes de um sistema complexo são heterogêneos, o que descreve os modelos baseados em agentes que tal autor estuda. Todavia, essa definição não cobre todos os casos possíveis de complexidade: em um autômato celular, como o Jogo da Vida, peças que estão sujeitas às mesmas regras podem dar origem a padrões complexos. Feita essa ressalva, podemos generalizar a definição.

“Mais precisamente, um fenômeno é emergente se ele exige novas categorias para descrevê-lo que não são exigidas para descrever o comportamento dos componentes subjacentes.” (GILBERT; TROITZSCH, 2005, p. 11, tradução própria). Para Miller e Page (2007, p. 78–91), a ferramenta adequada para estudar o comportamento de sistemas complexos seria a *simulação baseada em agentes* autônomos capazes de agir de maneira autônoma em um ambiente em busca de seus objetivos (SIEGFRIED, 2014, p. 18). As múltiplas possibilidades de ação dos agentes entre si e com o ambiente podem dar origem a comportamentos que não são previstos nas regras iniciais que os governam, dando origem assim a comportamentos emergentes.

4.1.4 Ficções e idealizações de modelagem

A construção de um modelo implementável como simulação envolve a tradução de conceitos teóricos para uma linguagem matemática tratável no computador, o que pode exigir definições alternativas de conceitos, expressões formais ou mesmo desvios em relação à teoria original. Além das modificações motivadas por restrições computacionais, a escassez de dados sobre um determinado fenômeno, ou a impossibilidade de observar mecanismos que seriam relevantes, pode levar à introdução no modelo de entidades, relações ou valores de parâmetros que sabidamente não possuem um referente no sistema-alvo¹². Em ambos os casos, estamos diante de *ficções*¹³ que deliberadamente relativizam o caráter mimético da simulação em relação ao sistema-alvo.

Para Weisberg (2013, p. 98–114), estaríamos diante de idealizações na modelagem, tanto para simplificação do modelo (WEISBERG, 2013, p. 99–100), quanto para identificação de seus fatores causais centrais (WEISBERG, 2013, p. 100–103). Uma terceira hipótese também foi mencionada por nós na Seção 2.4: a introdução de idealizações pode simplificar a integração de *idealizações de múltiplos modelos*, isto é, de sistemas que são formados por vários modelos que partem de hipóteses incompatíveis entre si. Neste caso, destaca Weisberg (2013, p. 103), não seria possível obter uma desidealização que desfaça as ficções estabelecidas e forneça uma representação “correta”, de forma que os modelos são a princípio irreconciliáveis entre si: no exemplo de simulação de fratura de placa metálica discutido na Seção 2.4 desta dissertação, temos três modelos diferentes do comportamento de fratura de placas, que partem de premissas incompatíveis para descrever regiões distintas do corpo estudado.

¹² Para Venturini, Pablo Jensen e Latour (2015) o advento das grandes bases de dados típicas de *big data* tornaria desnecessário o uso destas ficções e, em última instância, a prática simulacional em si, já que não seria necessário gerar dados artificiais em um cenário de abundância de dados. Argumentamos em sentido contrário na Seção 1.2, mas, ainda que se aceite o argumento de Venturini, Pablo Jensen e Latour (2015), a construção de simulações, inclusive fazendo uso de ficções, pode ser conveniente no caso prático.

¹³ Este uso de ficções como elemento de modelagem não deve ser confundido com o *ficcionalismo*, perspectiva teórica segundo a qual as simulações em si são cenários fictícios do sistema-alvo, tema que é discutido, dentre outros, por Weisberg (2013, p. 46–70).

A conciliação nas idealizações de múltiplos modelos exigirá *trade-offs* entre as diferentes abordagens, sacrificando elementos de uma ou de outra em prol dos objetivos globais do sistema. Esta harmonização, como ressalta Weisberg (2013, p. 105–110), pode ser guiada por diferentes ideais, que fornecerão critérios distintos de aceite. O primeiro seria a *completude* (WEISBERG, 2013, p. 105–106): a melhor representação de um sistema seria aquela que inclui todas as propriedades dele e todos os fatores externos que podem influenciá-lo, e as idealizações seriam reguladas pelo seu impacto neste objetivo último. Outro objetivo seria a *simplicidade* (WEISBERG, 2013, p. 106), na qual se busca a representação mínima que ainda satisfaça os critérios de fidelidade adotados.¹⁴ Os ideais admitem variações e aproximações que incorporem outras virtudes epistêmicas, mas servem de base para a discussão sobre o que relativizar ou não frente às demandas práticas envolvidas na construção de uma simulação.¹⁵

4.2 Coleta e uso de dados

O processo de construção de simulações computacionais é dependente dos dados que estão disponíveis sobre o fenômeno que se pretende estudar. Com base nestes dados é que se realizará a abstração do sistema-alvo — e, portanto, a escolha dos modelos — e a validação da simulação implementada para os objetivos propostos. Além disso, os dados de observação servirão de base para as análises a serem feitas sobre os resultados das simulações, permitindo o teste das hipóteses obtidas a partir do ambiente *in silico* em condições observadas fora da simulação. Esta dependência dos dados é relativizada em simulações com alvos abertos ou que se propõem a reproduzir cenários inteiramente fictícios, mas é o que permite o estabelecimento de representações para a explicação ou predição de fenômenos externos à simulação.¹⁶

A coleta dos dados pode ser feita manualmente, através de observadores humanos que codificam suas observações, ou de maneira automatizada (CIOFFI-REVILLA, 2014, p. 67–88). O uso de processos automáticos para capturar os dados e representá-los permite que a informação relevante seja produzida de maneira flexível, transparente para o usuário, coerente e ágil, aumentando o número e o tipo de eventos que podem ser capturados (BEIELER et al., 2016) e, portanto, empregados na construção da simulação. Estas técnicas, no entanto, introduzem um limitante técnico ao tipo de dado que pode ser

¹⁴ Embora o tratamento de abstrações feito por Weisberg (2013) seja típico da tradição filosófica dita *analítica*, isso não significa que as noções de idealização aqui discutidas não sejam relevantes para outras culturas dentro das ciências sociais. Como mostra Halas (2018), a noção marxiana de “abstração”, tida como pressuposto metodológico das pretensões de cientificidade de Marx e da tradição por ele influenciada, envolve uma omissão tácita de propriedades que é compatível com a ideia de idealização minimalista presente em Weisberg (2013).

¹⁵ Com base em um entendimento do papel instrumental das idealizações semelhante ao aqui proposto, Jebeile (2017) faz uma discussão mais extensa do uso da idealização como ferramenta do construtor de simulações, ressaltando as aplicações na transformação do modelo conceitual em termos matemáticos.

¹⁶ Quando ela possui um referente externo.

coletado, já que certas formas de informação — valores de métricas, informação básica de textos — são melhor tratadas computacionalmente que outras, como impressões pessoais.

Os dados sobre um fenômeno social, como discutido na Subseção 2.4.1, não são observações brutas de uma realidade externa, mas são condicionados pelos compromissos teóricos e valorativos dos construtores da simulação e dos responsáveis pela coleta desses dados. O escrutínio dos dados coletados, portanto, exige uma análise das premissas envolvidas na seleção das métricas relevantes, que devem capturar as entradas necessárias para o modelo construído, e na definição do que é considerado dado válido, já que as observações podem conter erros de vários tipos.

Considerações de ordem prática podem reduzir a capacidade de captação das métricas identificadas como relevantes. Por exemplo, em um cenário de marketing digital, pode ser interessante ter uma noção do consumo do usuário em outros *sites* para projetar a propensão de compra do cliente em potencial, mas a empresa que desenha o modelo pode não ter um meio efetivo de coletar esses dados, seja pela dificuldade de posicionar sensores, seja por restrições legais ou de negócio. Ainda que existam esses dados, é possível que eles não tenham a qualidade necessária, ou que a coleta na qualidade necessária não faça sentido dada as restrições de vista de tempo ou dinheiro.

A escassez de dados é um problema que pode dificultar a construção de simulações de um sistema, mas mesmo em um cenário no qual se dispõe de muita informação sobre o sistema-alvo, é possível que esta não esteja acessível de maneira mais imediata para a construção da simulação. Cai e Zhu (2015) indicam quatro dificuldades na coleta de dados em um ambiente de *big data*:

1. a integração de dados estruturados de formas complexas e vindos de várias fontes;
2. a dificuldade em avaliar em tempo hábil grandes volumes de dados;
3. o caráter perecível das informações, que podem perder rápido sua validade; e
4. a falta de padrões internacionais de qualidade de dados.

A combinação entre os quatro fatores pode dificultar a identificação, dentro de uma grande massa de dados, da informação relevante para a construção da simulação, problema que pode ser mitigado por técnicas de *data mining*.¹⁷ ou pelo uso exploratório da própria simulação, como discutido na Seção 3.2

Quando não há acesso a informações relevantes, a modelagem precisa encontrar métodos de compensar essa lacuna. Uma forma comum é o uso de estimativas, isto é, trabalhar os valores faltantes a partir de um valor fixo ou uma distribuição que representem

¹⁷ Alvarez (2016) traz algumas das abordagens possíveis em dados simulacionais.

o comportamento real da variável faltante.¹⁸ Outro mecanismo é a adoção de ficções, isto é, pressupostos sabidamente falsos que cumprem apenas um papel de integração do modelo, possibilitando seu funcionamento sem a pretensão de representar a realidade externa nesse ponto.

4.3 Implementação e execução de simulações

Simulações computacionais possuem modelos conceituais em seu cerne, mas estes modelos, para produzir os resultados desejados, passam por um processo de desenvolvimento que culmina em sua execução em um sistema computacional real. A escrita do código pode introduzir problemas de lógica ou de interpretação do modelo teórico, e as ferramentas empregadas podem conter *bugs* ou casos limítrofes inesperados, como descrito na Seção 1.3. Uma forma de evitar estes problemas é o uso de padrões de estilo de código (CIOFFI-REVILLA, 2014, p. 34–35), de ferramentas e de abordagens computacionais que já passaram por processos de verificação e validação (GILBERT; TROITZSCH, 2005, p. 21). Esta padronização facilita também o tratamento de ambiguidades conceituais, viabiliza o desenvolvimento colaborativo e estabelece objetivos de qualidade para as soluções (SIEGFRIED, 2014, p. 63–68).

Um modelo de referência, para Siegfried (2014, p. 69–75), deve distinguir claramente entre o modelo computacional e o programa que o implementa, ser capaz de lidar com diversos tipos de ambientes e de arquiteturas de agentes, lidar com ações, restrições e efeitos, e ser especificado de maneira formal. Embora tais requisitos sejam definidos por Siegfried (2014) para modelos baseados em agentes, os requisitos de formalização e de separação entre modelo conceitual, programa e arquitetura são aplicáveis a todas as abordagens de simulação, seguindo boas práticas de engenharia de software (SOMMERVILLE, 2011). Estes modelos são, então, instanciados para os problemas específicos, recebendo adições circunstanciais sobre a base original, que servirá como elemento para a construção da simulação final.

Em alguns casos, não existe uma solução pronta que seja adequada para a simulação que se deseja construir, o que torna necessário estender as funcionalidades de uma ferramenta existente ou mesmo criar uma solução nova. Neste caso, é interessante que a linguagem de programação a se utilizar atenda a certos requisitos (GILBERT; TROITZSCH, 2005, p. 21–22):

1. Permitir refinamentos incrementais do sistema;¹⁹

¹⁸ Um exemplo de técnica estatística para o tratamento de dados censurados ou faltantes é o algoritmo EM, discutido por Efron e Hastie (2016, p. 146–150).

¹⁹ Por exemplo, com o uso de interpretadores que permitam mudanças no código sem um processo extenso de compilação.

2. Gerar código portátil para múltiplas plataformas;
3. Executar os programas de maneira eficiente, permitindo múltiplas iterações de simulação para as análises de sensibilidade;²⁰
4. Ser familiar para o usuário e para a comunidade relevante, permitindo replicação da simulação;
5. Gerar código de depuração simples e rápida; e
6. Oferecer boas ferramentas e bibliotecas, em especial para visualização.

De Marchi (2005, p. 115–121) enfatiza a flexibilidade e a clareza como critérios de escolha das linguagens de programação e ambientes de desenvolvimento para a construção de uma simulação, bem como a modularização do código em funções, critérios que complementam a lista de Gilbert e Troitzsch (2005).

Uma simulação envolve não só o seu modelo conceitual e a implementação deste modelo, mas também a *infraestrutura de simulação*, definida por Siegfried (2014, p. 16, tradução própria) como “(...) a totalidade do *hardware* e *software* (sistema operacional e motor de simulação) necessários para executar um modelo de simulação.”²¹ Assim, os requisitos para o desenvolvimento da simulação devem ser analisados não só em termos das capacidades da linguagem de programação, mas à luz do tamanho da simulação desejada e do poder computacional disponível para sua execução.

Duas estratégias para aumentar os recursos computacionais disponíveis são o uso de *computação paralela* — isto é, a execução de uma mesma simulação usando várias *threads*, processadores ou nós de computação que compartilham recursos (SIEGFRIED, 2014, p. 50) —, ou de *computação distribuída*, em que a simulação é executada em múltiplos nós de computação que não compartilham recursos entre si (SIEGFRIED, 2014, p. 49–61). Em ambos os casos, o uso adequado dos recursos adicionais pode levar à redução do tempo de execução das simulações e facilitar a criação de mecanismos para a tolerância a falhas, bem como viabilizar a execução de *ensembles* de simulações autônomas (SIEGFRIED, 2014, p. 49). Todavia, este uso traz algumas dificuldades de ordem prática: passa a ser necessário garantir que os nós de computação consigam se comunicar uns com os outros, e que essa comunicação ocorra de forma síncrona e preservando a causalidade dos eventos.²²

²⁰ Este objetivo pode levar a uma preferência por linguagens compiladas, que deve ser ponderada com o primeiro requisito a partir das necessidades do caso prático.

²¹ No original: “(...) the entirety of hardware and software (operating system as well as simulation engine) necessary to execute a simulation model”.

²² Este problema é em parte mitigado nas simulações em tempo discreto, nas quais cada instante pode servir como unidade de referência para as operações de sincronização do sistema.

4.4 Calibração e validação de simulações

Uma vez que o modelo conceitual do fenômeno tenha sido convertido em um programa simulacional, ele pode em princípio ser executado em um sistema computacional concreto e produzir resultados. Para que estes resultados satisfaçam os objetivos desejados para a simulação, é necessário *calibrar* os parâmetros de execução de maneira satisfatória. Além disso, caso este objetivo envolva a execução continuada da simulação ou a realização de múltiplas execuções, pode ser necessária a *manutenção* do programa implementado, de forma a ajustá-lo para lidar com mudanças das circunstâncias da simulação, do seus casos de uso ou do *hardware* em que ela é executada.

Para tentar compensar vieses de confirmação, simulações mais complexas fazem uso tanto das construções teóricas quanto de dados empíricos para a validação de seus resultados (VENTURINI; JENSEN, P.; LATOUR, 2015). Os dados coletados a partir da observação do sistema-alvo podem desempenhar múltiplos papéis no desenho de uma simulação.

Um primeiro uso envolve o desenho dos cenários de simulação: os dados reais podem ser utilizados para definir o estado inicial da simulação — por exemplo, a população inicial de um modelo baseado em agentes e sua distribuição —, seja através da replicação direta dos estados observados, seja através do estabelecimento de cenários que variem um ou mais aspectos em relação a um cenário realmente observado. Outro uso semelhante pode ser identificado na determinação dos parâmetros que controlam o modelo subjacente à simulação, cujos valores podem ser diretamente extraídos através de abordagens analíticas, ou inferidos estatisticamente a partir das variáveis observadas.²³

Um terceiro modo de uso dos dados empíricos na calibração do modelo é a comparação das saídas da simulação implementada com os estados observados do sistema-alvo da simulação. A execução do programa simulacional partindo de um cenário suficientemente parecido com um cenário real deveria²⁴ produzir saídas semelhantes às verificadas no caso concreto.²⁵ Embora a escolha das métricas para determinar esta similaridade seja dependente do cenário específico de modelagem,²⁶ a avaliação destas métricas pode ser

²³ Os parâmetros podem, em uma abordagem frequentista, ser computados como uma estatística representativa da população de observações do sistema real ou, sob uma perspectiva bayesiana, serem eles mesmo variáveis aleatórias. A diferença entre ambas as abordagens não é significativa para a presente análise das dificuldades de validação, mas dá origem a tratamentos radicalmente diferentes no uso de métodos estatísticos para a construção da simulação. Efron e Hastie (2016) apresentam um panorama do estado da arte em ambas as vertentes de inferência estatística.

²⁴ Uma exceção a este princípio pode ser identificada no caso dos sistemas caóticos, isto é, sistemas dinâmicos cujas saídas são altamente sensíveis às condições de entrada. Nesse caso, é possível que pequenas diferenças em parâmetros da simulação levem a saídas radicalmente diferentes.

²⁵ Em um sistema puramente determinístico, a saída de uma execução da simulação seria o objeto comparável ao sistema-alvo. Já em simulações com componentes probabilísticos, serão necessárias várias execuções para estabelecer uma comparação entre as propriedades estatísticas da distribuição das saídas da simulação e os dados observados no caso concreto.

²⁶ Em uma simulação de agentes voltada ao estudo de desigualdades de renda, por exemplo, métricas

feita a partir da aplicação de técnicas de inferência estatística, como as descritas em [Efron e Hastie \(2016\)](#), para obter as médias e distribuições das variáveis observadas nas saídas da simulação e compará-las com os valores observados no sistema-alvo.

Quando falamos em calibração dos parâmetros, podemos usar o termo em um sentido estrito — como faz [Weisberg \(2013, p. 93–94\)](#) ao definir um modelo calibrado como uma representação de alta fidelidade do sistema-alvo, inclusive em termos quantitativos — e um sentido mais amplo. No sentido mais estrito, a fidelidade da representação pode ser estabelecida a partir da análise estatística dos resultados da simulação para múltiplos valores dos parâmetros que precisam ser determinados, comparando os resultados com os dados observados do sistema a partir de testes de hipótese ou outras técnicas existentes. Tal processo de calibração é essencial para a validação das inferências extraídas desse modelo, ao estabelecê-lo como um instrumento adequado para a observação, ainda que indireta, do sistema-alvo,²⁷ evitando problemas como o sobreajuste²⁸ do modelo aos dados.

No entendimento mais amplo da noção de calibração, se deseja verificar o ajuste *qualitativo* dos resultados da simulação ao comportamento conhecido do sistema-alvo. Aqui a análise estatística é menos útil em um primeiro momento, mas ainda é possível estabelecer critérios de avaliação, que também podem ser aplicados à calibração quantitativa. [Cioffi-Revilla \(2014, p. 238–243\)](#) traz três critérios para a avaliação de modelos formais em ciências sociais: a *verdade*, entendida como o poder de explicação do modelo;²⁹ a *beleza*, definida em termos de propriedades como o estilo adotado na formalização e a economia nas estruturas adotadas; e a *justiça*, vista como o impacto do modelo na construção de um mundo melhor. Estes critérios³⁰ podem ser relativizados para certos tipos de simulação — em simulações inteiramente fictícias, seria mais adequado falar em *verossimilhança* do que em verdade — e devem ser complementados pela análise dos requisitos computacionais, como o tempo de execução e itens discutidos nas seções anteriores, mas fornecem diretrizes para a escolha dos parâmetros de execução de uma simulação.

relevantes são aquelas que capturam a renda dos indivíduos e as diferenças dentro da população. Já em um cenário de segregação como o de [Schelling \(1972\)](#), descrito no Capítulo 1, seriam avaliadas métricas de concentração e dispersão populacional.

²⁷ Sobre o papel da instrumentação na observação indireta de sistemas, a discussão presente em [Humphreys \(2004, cap. 2\)](#) pode suplementar os argumentos de adequação apresentados nesta dissertação.

²⁸ É comum, mesmo na literatura em língua portuguesa, o uso da versão anglófona *overfitting*.

²⁹ [Hasse e Lenhard \(2017\)](#) esclarecem um duplo papel dos parâmetros na epistemologia dos modelos, que pode ser vista dentro do critério de *verdade* aqui posto: por um lado, o processo de ajuste dos parâmetros pode permitir que uma simulação reproduza seu objeto de maneira mais precisa que o esperado com base no conhecimento teórico disponível; por outro, o ajuste de parâmetros pode comprometer o valor epistêmico da simulação ao introduzir fatores estranhos, por exemplo em parâmetros de implementação sem correspondentes em um modelo teórico do fenômeno.

³⁰ A determinação do conteúdo geral destes três valores está ligada a problemas de longa história na filosofia, mas técnicas de projeto de *software* como o design sensível a valores e o design participativo podem ser utilizados para propor como a beleza, a justiça e a verdade devem ser interpretadas no caso concreto, a partir da consideração das múltiplas perspectivas de usuários e das pessoas direta e indiretamente afetadas pela simulação ([ALMADA; ATTUX, 2018](#)).

Existem, no entanto, problemas de ordem teórica que devem ser superados por qualquer processo de avaliação estatística das saídas de uma simulação. Quando os dados são empregados na definição do cenário da simulação e de seus parâmetros, o modelo subjacente é ajustado para melhor descrever estes dados, o que pode levar a um sobreajuste do modelo.³¹

Para mitigar o problema de sobreajuste, as soluções típicas da literatura³² fazem uso de técnicas de validação cruzada, em que os parâmetros de um modelo são avaliados a partir da segmentação dos dados disponíveis, de forma que todos os dados disponíveis sejam usados no treinamento e na validação de um modelo, mas em momentos distintos; isto é, um dado nunca é ao mesmo tempo parte do conjunto de treinamento e do conjunto de validação (EFRON; HASTIE, 2016). Esta abordagem é particularmente interessante para a validação de parâmetros de simulações, ao permitir que a construção do modelo se beneficie de toda a informação disponível, e é implementada por padrão em várias bibliotecas e ferramentas estatísticas.

Não é nosso intuito nesta seção explorar exaustivamente as abordagens de verificação e validação de sistemas computacionais, mas é importante ressaltar as aplicações de boas práticas deste campo no domínio das simulações computacionais. Sargent (2013) apresenta um paradigma de construção de simulações em que a verificação e a validação da simulação são entrelaçadas, ao longo de um conjunto de etapas de desenvolvimento, com diferentes conceitos de validade do modelo subjacente. Já Ngo e See (2012) exploram um processo passo-a-passo: inicialmente se faz uma verificação de plausibilidade dos resultados iniciais, seguida por uma análise de sensibilidade dos parâmetros do modelo; esta verificação inicial é seguida por uma calibração dos parâmetros para obter a melhor descrição do sistema-alvo; e a última fase consiste na validação das saídas do modelo. Estas e outras abordagens de verificação e validação de simulações computacionais buscam estabelecer uma correspondência entre o alvo da simulação e os dados simulados, permitindo que estes sejam objeto de análise científica.

O acompanhamento da adequação dos resultados de uma simulação não deve ficar restrito à sua etapa de desenvolvimento. Mesmo que a validade da simulação tenha sido estabelecida em um momento inicial, é possível que ela se torne inválida em virtude de mudanças no cenário de execução — por exemplo, que façam o modelo atuar em uma região para a qual ele não foi validado inicialmente ou pela mudança nos tipos ou fontes de dados que alimentam a simulação —, por erros de execução que só possam ser detectados ao longo do uso continuado do programa simulacional, por problemas na infraestrutura de simulação ou mesmo porque o objetivo de uso da simulação foi alterado em relação ao momento de sua criação. Como todos estes cenários podem ocorrer, especialmente para

³¹ Um modelo sofre sobreajuste quando sacrifica parte de sua capacidade de generalização em prol de um melhor ajuste aos dados do conjunto de treinamento.

³² Ver, por exemplo, Provost e Fawcett (2013).

as simulações que são executadas em fluxo contínuo, os programas simulacionais devem ser validados constantemente, através da realização de testes automatizados e outras técnicas de engenharia de *software*³³ que garantam a validação das saídas da simulação e a calibração de seus parâmetros para o fim desejado.

4.5 Análise e uso dos resultados

A construção, a calibração e a validação das simulações computacionais são feitas, em última instância, para justificar os usos que se faz dos resultados de sua execução. Seja a simulação um ponto de partida para a análise do fenômeno modelado, uma ferramenta de predição ou um mero exercício de ficção — dentre outras finalidades discutidas no Capítulo 3 —, as etapas anteriores, se bem executadas, garantem que ela produzirá saídas computacionais que sejam adequadas às finalidades de sua criação. Estas saídas serão analisadas e interpretadas com base no modelo da simulação, de forma que os resultados computacionais possam servir de base para inferências sobre o fenômeno simulado.

As saídas de uma simulação computacional, assim como suas entradas, representam o sistema-alvo de uma forma codificada (DE MARCHI, 2005, p. 121). Uma vez que esta codificação é conhecida pelo processo de construção da simulação, os dados gerados podem ser entendidos a partir dos isomorfismos parciais entre o modelo subjacente à simulação e a descrição do alvo, invertendo o processo descrito na Seção 2.4, o que torna viável o uso dos métodos das ciências sociais adequados para o problema.

A escolha dos métodos adequados será sempre dependente do objeto de estudo da simulação e da finalidade para a qual ela é usada: uma simulação de *churn* de clientes de um serviço bancário será avaliada a partir de métricas de prejuízo financeiro com a perda de clientes, por exemplo, ao passo que uma implementação do modelo de segregação de Schelling (1972) seria avaliada de uma perspectiva qualitativa. Existem, no entanto, algumas características gerais das abordagens de análise e dos usos dos resultados de simulações, que discutiremos nesta seção.

Weisberg (2013, p. 79–83) faz uma distinção entre *análises completas* dos resultados de uma simulação, em que o objetivo é representar o estado total do modelo, e *análises parciais*, em que se busca responder a uma questão mais específica, como o impacto de uma alteração em um de seus parâmetros. Em ambos os casos, o processo de análise envolverá a calibração do modelo (WEISBERG, 2013, p. 90) e, para simulações estocásticas, a realização de múltiplas execuções com os mesmos parâmetros, de forma a elucidar o comportamento estatístico do modelo sob as mesmas condições de execução e evitar que se tirem conclusões a partir de casos pouco representativos caso não seja essa

³³ E.g., Sommerville (2011) e David (2013).

a intenção.³⁴

Essa repetição das execuções do modelo e da subsequente análise dos resultados tem como objetivo principal garantir a *robustez* das inferências feitas sobre o sistema-alvo com base na simulação. A robustez dos resultados de uma simulação, já discutida na Seção 1.3 a partir da noção de simulação robusta de Muldoon (2007), contribui para a confiança nas inferências neles baseadas ao garantir que estas se sustentam frente a mudanças nas condições do problema, servindo de base para a *generalização* destas afirmações.

Uma primeira forma de robustez dos resultados é a *robustez nos parâmetros* (WEISBERG, 2013, p. 160–161), em que uma propriedade se conserva com a mudança de parâmetros, o que pode ser investigado diretamente por exaustão ou por técnicas de amostragem do espaço de parâmetros.³⁵ Há também a *robustez estrutural* (WEISBERG, 2013, p. 161–162), que percebe o impacto das mudanças nos aspectos mecânicos do sistema, como as regras de causalidade nos resultados obtidos. Por fim, Weisberg (2013, p. 162–163) fala em *robustez de representação*, que conserva os parâmetros e aspectos mecânicos, mas vê como a mudança de representação pode afetar um sistema, por exemplo, através da troca de um modelo baseado em população por um modelo baseado em indivíduos.

A análise dos resultados pode ser o objetivo direto da simulação, ao servir de base para a formação de teorias sobre o sistema-alvo, ou ser simplesmente um estágio de validação externa da simulação, estabelecendo a qualidade das saídas e das inferências derivadas que serão utilizadas por outro processo computacional ou para sustentar a tomada de decisão de atores sociais. De uma forma ou de outra, a robustez dos resultados contribui para a confiabilidade dos resultados, que podem então ser sujeitos a testes de replicação ou outros mecanismos para assegurar a *accountability* no emprego das simulações.

4.6 Limites não-técnicos das simulações

Como discutido nos capítulos 1 e 3, simulações são objetos que podem ser utilizados como fonte de sustentação de discursos sobre domínios sociais, como o planejamento de políticas públicas (DESAI, 2012). Tendo em vista os impactos possíveis destes discursos na vida de indivíduos e grupos sociais e políticos, o uso de simulações computacionais está sujeito a diversas restrições sociais, bem como sanções para consequências socialmente indesejáveis, como a morte de indivíduos em decorrência de decisões fundamentadas por

³⁴ A realização de múltiplas execuções e de análises potencialmente complexas sobre os resultados de cada iteração pode ser simplificada, até certo ponto, pela automação dos processos de calibragem e execução da simulação e de processamento dos resultados, que será de fato essencial para problemas de grande porte.

³⁵ Gilbert e Troitzsch (2005, p. 25) destacam o papel da aleatorização dos parâmetros como forma de obter uma amostragem razoável do espaço de variáveis e fazer análise de sensibilidade dos resultados.

uma simulação (HALLEVY, 2010).³⁶

Na Seção 3.3, tratamos do impacto da inteligibilidade de uma simulação na aceitação ou na rejeição de argumentos nela baseados. Ali, a discussão proposta focou na questão dimensão argumentativa, mas esta pode ter impactos legais, na medida em que simulações, em especial aquelas envolvendo alguma forma de inteligência artificial, são utilizadas em contextos nos quais existem exigências legais de que os resultados possam ser explicados (DOSHI-VELEZ et al., 2017), como é o caso das decisões judiciais no sistema jurídico brasileiro.

Imposições e permissões de origem legal podem entrar em choque com requisitos éticos presentes no uso da simulação. Um exemplo é o estudo de redes sociais: embora os termos de uso de *sites* como o Twitter permitam a coleta de dados dentro de certas condições, mesmo estes dados abertos podem ser utilizados de forma a revelar a identidade de ativistas, como descreve Morozov (2012). Uma simulação que se baseie neste tipo de dados pode dar origem a problemas éticos, em especial em cenários ditatoriais, ainda que seja construída dentro dos termos da lei.³⁷

É necessário, como sugere Alvarez (2016), preparar pesquisadores e comissões de ética para lidarem com os problemas éticos e legais que podem surgir das pesquisas sociais envolvendo grandes volumes de dados. Em usos fora da academia, há também a necessidade de lidar com exigências de *compliance* e outras ferramentas de controle econômico, político e social que estabeleçam limites às atividades de pesquisa sobre o comportamento humano. Através desses controles e cuidados ao longo do desenvolvimento e uso das simulações, é possível realizar o objetivo de justiça trazido por Cioffi-Revilla (2014, p. 238–243), voltando estes modelos para melhorar a condição humana.

³⁶ No sistema jurídico brasileiro, já existe literatura a respeito da responsabilidade civil em virtude de ações de sistemas computacionais, como Lopes e Roberto (2018) e Almada (2018), embora o foco da discussão esteja em sistemas autônomos e aprendizado de máquina.

³⁷ Uma discussão mais extensa dos aspectos éticos envolvidos no ciclo de desenvolvimento de simulações pode ser vista em Almada e Attux (2018).

5 Conclusão

— *As cidades também acreditam ser obra da mente ou do acaso, mas nem um nem outro bastam para sustentar as suas muralhas. De uma cidade, não aproveitamos as suas sete ou setenta e sete maravilhas, mas a resposta que dá às nossas perguntas.*

— *Ou as perguntas que nos colocamos para nos obrigar a responder, como Tebas na boca da Esfinge.*

Italo Calvino, *As Cidades Invisíveis*: As cidades e os mortos 1

O objetivo desta dissertação foi identificar as condições de contorno para que as simulações computacionais possam contribuir para o estudo — acadêmico ou não — de fenômenos sociais, bem como os efeitos práticos que simulações podem produzir na sociedade. Para isto, os capítulos anteriores extraíram uma noção de simulação a partir de usos já existentes de simulações em domínios sociais e do ferramental teórico fornecido pelas literaturas de filosofia da ciência e engenharia de *software*. A partir de tal delimitação, discutimos como as simulações poderiam ser ferramentas úteis para diversas abordagens de explicação científica e também para usos não necessariamente explicativos, como a predição de eventos futuros ou a construção de ficções didáticas ou lúdicas.

Uma peculiaridade do tratamento adotado nesta dissertação é considerar as simulações computacionais a partir de duas perspectivas distintas. De um lado, os tratamentos teóricos das simulações computacionais, baseados principalmente em técnicas simulacionais de Física e das engenharias, costumam analisá-las enquanto *objetos lógico-conceituais*, enfatizando o seu papel epistêmico na prática científica. De outro, a literatura dedicada aos impactos sociais da tecnologia costuma tratar as simulações como sistemas capazes de produzir efeitos diretos e indiretos no mundo exterior aos sistemas computacionais que as executam. No entanto, perceber estas duas visões como categorias-estaque pode ser contraproducente nos domínios de nosso interesse, já que as próprias simulações podem produzir efeitos no fenômeno social que elas se propõem a descrever¹ e a complexidade envolvida na maior parte destas aplicações exige mudanças no entendimento dos mecanismos de modelagem e implementação para que elas atinjam os objetivos desejados. Assim, a construção e o uso de simulações computacionais de fenômenos sociais deve ter em conta tanto o aspecto lógico-conceitual quanto o impacto social destes objetos.

¹ Por exemplo, simulações de comportamento do mercado financeiro podem influenciar atores que nele operam, alterando as dinâmicas que a simulação pretende reproduzir.

A discussão sobre a dimensão lógico-conceitual das simulações computacionais está centrada no Capítulo 2 desta dissertação, em que investigamos os papéis que elas podem desempenhar na produção de conhecimento — em particular, conhecimento científico. Esta investigação, amparada em linhas gerais no trabalho de [Frigg e Reiss \(2009\)](#), situou as simulações computacionais em relação a teorias e problemas filosóficos já amplamente trabalhados pela literatura, operando algumas reduções de características salientes das simulações.

Uma primeira redução operada está relacionada à natureza da atividade simulacional, já que parte da literatura² trata as simulações como uma ferramenta diferente tanto dos experimentos físicos quanto dos experimentos mentais. Em um primeiro momento, a criação de uma nova categoria parece necessária, já que a reconstrução que fizemos dos argumentos de [Beisbart e Norton \(2012\)](#) evidencia que muitas das formas simulacionais aceitas — como os métodos de Monte Carlo — não podem ser percebidas como experimentos nos moldes tradicionais, ao passo que enquadrar as simulações como experimentos mentais seria problemático por desconsiderar inteiramente o impacto epistêmico da implementação das simulações em sistemas computacionais reais e por alterar a forma como a relação entre a simulação e seu usuário é entendida.

Nenhum destes obstáculos à classificação das simulações como uma forma de experimento mental é, contudo, intransponível. A noção de que as simulações podem ser fontes de conhecimento — e não meramente argumentos — torna-se sustentável a partir da ênfase na simulação enquanto fonte de *informações acionáveis* ([GRECO, 2015](#)) sobre um domínio: assim, mesmo que as simulações estejam sujeitas a uma forma de fecho epistêmico, elas ainda assim podem tornar acessível um conhecimento que, apesar de logicamente disponível, não seja evidente antes do uso simulacional. Já as diferenças entre um experimento mental tradicional e um programa de computador que se proponha ao mesmo papel — como a possibilidade de falhas de *hardware* e a opacidade das etapas de execução — podem ser mitigadas ou até mesmo evitadas com técnicas de engenharia de *software* como as discutidas no Capítulo 4. Com estes cuidados de desenvolvimento, as simulações podem ser usadas como *amplificadoras* de experimentos mentais: através da implementação computacional, o investigador pode realizar *Gedankenexperimente* envolvendo um número maior de objetos e relações do que seria possível sustentar na cognição humana não-assistida, o que por sua vez torna viável a modelagem de fenômenos complexos, como a integração de múltiplos níveis de modelagem discutida na Subseção 4.1.4.

A literatura sobre simulações computacionais costuma enfatizar seu uso *explicativo*, isto é, o emprego de uma simulação na busca por fatores que indiquem a ocorrência ou não de um dado fenômeno. Nas ciências naturais, esta explicação em geral é associada à

² [Saam \(2017b\)](#) oferece um panorama da discussão, que serviu de base para a discussão realizada no Capítulo 2.

busca de fatores causais; já em domínios sociais, outros modos de explicação, como a busca por narrativas ou pelas razões que levam a uma ação, são preferidos em determinados contextos, uma vez que a explicação causal é tida como insuficiente para capturar fenômenos sociais “sutis” (VENTURINI; JENSEN, P.; LATOUR, 2015). No entanto, como discutimos nos capítulos 1 e 3, as simulações computacionais podem desempenhar papéis investigativos mesmo em modelos não-causais de explicação, contribuindo para a investigação científica.

Os usos científicos das simulações computacionais, além de serem valorados por sua produção de conhecimento, também servem de base para outras aplicações, variando do lúdico à tomada de decisões estratégicas. Estas simulações são criadas dentro de sistemas conceituais e ideológicos específicos, como discutido no Capítulo 3, mas seus efeitos diretos e indiretos no mundo podem ser ressignificados pelos usos que diversos *stakeholders* fazem dos programas, de suas saídas e das interpretações destas.

Desta forma, seria insuficiente discutir as simulações computacionais de fenômenos sociais apenas pela perspectiva epistemológica. Por outro lado, enfatizar a discussão sobre os impactos sociais das simulações traria o risco de ignorarmos que estes efeitos decorrem de usos que se fundamentam na ideia de que as simulações são, de forma geral, uma ferramenta válida para a produção de conhecimento com pretensões de validade científica. Assim, optamos nesta dissertação por adotar uma visão de simulação que reconcilie estas duas imagens distintas sobre o uso e os significados destes objetos computacionais que pretendem descrever aspectos da sociedade.

5.1 Uma imagem das simulações computacionais

Para delimitar a visão lógico-conceitual das simulações de fenômenos sociais, fizemos uso de uma ampla literatura voltada à discussão do papel das simulações computacionais nas ciências naturais. Estas técnicas são parte já estabelecida da metodologia científica, mas sua discussão teórica ainda não produziu uma visão coerente das simulações, uma vez que as diferentes áreas que as estudam enfatizam questões distintas que ainda não são respondidas de uma maneira unificada (RESCH, 2017).

Resch (2017, p. 31, tradução própria) destaca quatro questões que devem ser esclarecidas por uma teoria coerente das simulações:

1. Como o conceito de simulação se encaixa em nosso entendimento do ser humano e de sua capacidade de adquirir conhecimento?
2. Que conhecimento em potencial a simulação oferece em campos que não podem ser primariamente descritos através de fórmulas matemáticas?

3. Como as ciências sociais e as humanidades contribuem para o entendimento do fenômeno da simulação?
4. O que as simulações podem contribuir para a solução de problemas sociais e políticos?

Responder de maneira satisfatória às quatro questões está além do escopo desta dissertação, mas a justificativa do uso das simulações em domínios sociais exigiu que abordássemos alguns dos temas destacados por [Resch \(2017\)](#). Na seção anterior, resumimos a discussão sobre o uso instrumental das simulações na aquisição e organização do conhecimento, bem como sua contribuição para campos menos amados a tratamentos quantitativos ou nos quais prevalecem modos não-causais de explicação. Também à quarta pergunta oferecemos uma resposta instrumental: as simulações contribuem na medida em que estes sistemas sociotécnicos se encaixam em contextos de uso propícios, seja na investigação conceitual, seja na aplicação prática, ao expandir as capacidades analíticas e decisórias dos *stakeholders* de um determinado cenário. Esta expansão pode ser uma fonte de novas questões sociais — como a discriminação baseada em algoritmos — mas pode também contribuir para mitigá-los, através da adoção de critérios de desenho como os requisitos de *verdade, beleza e justiça* discutidos no Capítulo 4.

Como destaca [Resch \(2017, p. 30\)](#), tal discussão não está no centro das atenções da pesquisa em ciências sociais, mas é crítica em tempos nos quais as simulações adquirem maior importância social. Não cabe a esta dissertação uma discussão mais ampla a respeito dos caminhos a serem tomados pela pesquisa que a terceira questão destaca, a ser conduzida por equipes interdisciplinares guiadas por cientistas sociais, mas vale destacar aqui a possibilidade de que estes estudos façam uso de técnicas simulacionais para analisar suas hipóteses, nas linhas da aplicação que [Thagard \(1993\)](#) faz para a filosofia da ciência.

5.2 Limites e possibilidades das simulações

Simulações computacionais de fenômenos sociais, é importante lembrar, já são empregadas em diversos usos explicativos e não-explicativos, como descrito no Capítulo 3. Estas aplicações servem de ponto de partida para identificar elementos que caracterizam as simulações, as condições de contorno e os limites existentes à inferência baseada nelas. Além disso, a discussão teórica abre novas avenidas de emprego destas técnicas, como a produção de explicações de fenômenos que não fiquem restritas à causalidade, mas permitam, por exemplo, a identificação das razões que motivam a conduta de um ator social ou a validação de narrativas explicativas.

Com base na caracterização das simulações, retomada nas seções anteriores, podemos identificar alguns dos limites ao seu uso. Uma primeira fonte de limitantes são os modelos teóricos utilizados: não só eles determinam quais são os objetos e relações a modelar, como também condicionam a própria definição de quais são os dados relevantes para uma simulação. Estes dados também trazem limites por conta própria à atividade simulacional, na medida em que dados pouco representativos ou coletados de maneira inadequada prejudicam a capacidade de generalização de simulações que os usam. A necessidade de executar as simulações em um sistema computacional realmente existente também traz limites, como o tempo de execução da simulação e suas demandas de recursos operacionais. Todos estes fatores afetam os resultados de um programa simulacional e, por consequência, as interpretações possíveis a partir deles, o que torna relevantes controles como os discutidos no Capítulo 4.

A identificação de limites teóricos e práticos à prática das simulações computacionais é útil para a validação de simulações nas linhas das aplicações discutidas, mas também sugere outras modalidades de uso além da construção de explicações com base científica sobre um dado fenômeno. O Capítulo 3 apresenta alguns destes outros usos possíveis das simulações, como a construção de jogos de computador, a realização de predições sobre eventos futuros ou a exploração das consequências de um corpo teórico. Estes usos introduzem requisitos específicos para que uma simulação produza os resultados desejados em sua construção: uma simulação que tenha fins didáticos deve ser compreensível para seu público-alvo, enquanto um programa usado no setor financeiro terá exigências específicas de *compliance*, e assim por diante. Desta forma, o efeito das simulações não é determinado apenas por suas propriedades epistêmicas, mas deve também levar em conta os objetivos e interesses de múltiplos *stakeholders*.

Tal multiplicidade de *stakeholders* também é relevante para a construção de explicações nas ciências sociais, tanto pelos interesses envolvidos na pesquisa científica quanto pela pluralidade da prática científica. O processo de construção das simulações explicita conceitos e relações implícitos em sistemas teóricos, facilitando o processo de tradução que Kuhn (2012, p. 197–203) destaca como necessário para a comunicação entre adeptos de paradigmas distintos. Ao fazê-lo, a simulação não só dá novas técnicas ao cientista, mas facilita a colaboração entre pesquisadores de várias áreas e entre campos teóricos distintos.

Bibliografia

- ALMADA, Marco. Responsabilidade civil extracontratual e a inteligência artificial. *Revista Acadêmica Arcadas*, 2018. A publicar.
- ALMADA, Marco; ATTUX, Romis. Ethical design of social simulations. In: *3º WASHES — Workshop sobre aspectos sociais, humanos e econômicos de Software*. Natal: Sociedade Brasileira de Computação, 2018. p. 17–26.
- ALVAREZ, R Michael (Ed.). *Computational Social Science: Discovery and Prediction*. Cambridge University Press, 2016.
- AXELROD, R. *The Evolution of Cooperation*. Basic Books, 1984.
- BEIELER, John et al. Generating political event data in near real time: Opportunities and challenges. In: *Computational Social Science*. Cambridge University Press, 2016.
- BEISBART, Claus. Advancing Knowledge Through Computer Simulations?: A Socratic Exercise. In: *The Science and Art of Simulation I*. Springer, 2017. p. 153–174.
- BEISBART, Claus; NORTON, John D. Why Monte Carlo simulations are inferences and not experiments. *International Studies in the Philosophy of Science*, Taylor & Francis, v. 26, n. 4, p. 403–422, 2012.
- BERLIN, Isaiah. The divorce between the sciences and the humanities. *Salmagundi*, JSTOR, n. 27, p. 9–39, 1974.
- BIANCHI, Federico; SQUAZZONI, Flaminio. Agent-based models in sociology. *Wiley Interdisciplinary Reviews: Computational Statistics*, Wiley Online Library, v. 7, n. 4, p. 284–306, 2015.
- BRACHMAN, Ronald; LEVESQUE, Hector. *Knowledge Representation and Reasoning*. San Francisco: Morgan Kaufmann Publishers Inc., 2004.
- BREWER, William F; LAMBERT, Bruce L. The theory-ladenness of observation and the theory-ladenness of the rest of the scientific process. *Philosophy of Science*, University of Chicago Press, v. 68, S3, s176–s186, 2001.
- CAI, Li; ZHU, Yangyong. The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, Ubiquity Press, v. 14, 2015.
- CARNOTA, Raúl; RODRÍGUEZ, Ricardo. AGM Theory and artificial intelligence. In: *Belief revision meets philosophy of science*. Springer, 2010. p. 1–42.
- CATTELL, Rick. Scalable SQL and NoSQL data stores. *ACM SIGMOD Record*, ACM, v. 39, n. 4, p. 12–27, 2011.

- CIOFFI-REVILLA, Claudio. *Introduction to computational social science*. New York: Springer, 2014.
- COGLIANO, Jonathan F; JIANG, Xiao. Agent-based computational economics: simulation tools for heterodox research. In: LEE, Frederick S; CRONIN, Bruce (Ed.). *Handbook of Research Methods and Applications in Heterodox Economics*. Cheltenham: Edward Elgar Publishing, 2016. cap. 13, p. 253–.
- CONTE, Rosaria et al. Manifesto of computational social science. *The European Physical Journal Special Topics*, Springer, v. 214, n. 1, p. 325–346, 2012.
- DAVID, Nuno. Validating simulations. In: *Simulating Social Complexity*. Heidelberg: Springer, 2013. p. 135–171.
- DAVIS, Janet. Design methods for ethical persuasive computing. In: ACM. *Proceedings of the 4th International Conference on Persuasive Technology*. 2009. p. 6.
- DAVIS, Paul K. *Aggregation, Disaggregation, and the 3:1 Rule in Ground Combat*. Santa Monica, 1995.
- DE MARCHI, Scott. *Computational and mathematical modeling in the social sciences*. Cambridge: Cambridge University Press, 2005.
- DESAI, Anand (Ed.). *Simulation for policy inquiry*. Springer Science & Business Media, 2012.
- DI PAOLO, Ezequiel A; NOBLE, Jason; BULLOCK, Seth. Simulation models as opaque thought experiments. In: *Seventh International Conference on Artificial Life*. Cambridge, MA: MIT Press, 2000. p. 497–506.
- DOSHI-VELEZ, Finale et al. Accountability of AI under the law: The role of explanation. *arXiv preprint*, 2017. arXiv:1711.01134.
- EFRON, Bradley; HASTIE, Trevor. *Computer age statistical inference*. Cambridge University Press, 2016. v. 5.
- ELGIN, Catherine Z. *True enough*. MIT Press, 2017.
- ELSENBROICH, Corinna. Explanation in agent-based modelling: Functions, causality or mechanisms? *Journal of Artificial Societies and Social Simulation*, v. 15, n. 3, 2012.
- ELSENBROICH, Corinna; GILBERT, Nigel. *Modelling Norms*. New York: Springer Publishing, 2013.
- ELSTER, Jon. *Logic and Society: Contradictions and Possible Worlds*. Chichester: John Wiley & Sons, 1978.
- ELTON, Geoffrey Rudolph. Two kinds of history. In: FOGEL, Robert William; ELTON, Geoffrey Rudolph (Ed.). *Which Road to the Past? Two Views of History*. New Haven: Yale University Press, 1983. p. 71–122.

- FISCHER, Bob. *Modal justification via theories*. Springer, 2017.
- FLORIDI, Luciano. On the Logical Unsolvability of the Gettier Problem. *Synthese*, Springer, v. 142, n. 1, p. 61–79, 2004.
- FOGEL, Robert William. “Scientific” history and traditional history. In: *Studies in Logic and the Foundations of Mathematics*. Elsevier, 1982. v. 104. p. 15–61.
- FOURCADE, Marion; HEALY, Kieran. Seeing like a market. *Socio-Economic Review*, Oxford University Press, v. 15, n. 1, p. 9–29, 2016.
- FREDKIN, Edward. An introduction to digital philosophy. *International Journal of Theoretical Physics*, Springer, v. 42, n. 2, p. 189–247, 2003.
- FRIGG, Roman; REISS, Julian. The philosophy of simulation: hot new issues or same old stew? *Synthese*, Springer, v. 169, n. 3, p. 593–613, 2009.
- GAGNEBIN, Jeanne Marie. Uma filosofia do *cogito* ferido: Paul Ricoeur. *Estudos Avançados*, scielo, v. 11, p. 261–272, ago. 1997.
- GAVIRA, Muriel de Oliveira. *Simulação computacional como uma ferramenta de aquisição de conhecimento*. 2003. Dissertação (Mestrado em Engenharia de Produção) – Universidade de São Paulo, São Carlos.
- GELFERT, Axel. *How to do science with models: A philosophical primer*. Springer, 2016.
- GELMAN, Andrew; KING, Gary. Enhancing democracy through legislative redistricting. *American Political Science Review*, Cambridge University Press, v. 88, n. 3, p. 541–559, 1994.
- GENDLER, Tamar Szabó. *Thought experiment: On the powers and limits of imaginary cases*. New York: Routledge, 2000.
- GETTIER, Edmund L. Is justified true belief knowledge? *Analysis*, JSTOR, v. 23, n. 6, p. 121–123, 1963.
- GHINS, Michel. Bas van Fraassen on Success and Adequacy in Representing and Modelling. In: *Model-Based Reasoning in Science and Technology*. Heidelberg: Springer, 2016. p. 21–42.
- GILBERT, Nigel; TROITZSCH, Klaus. *Simulation for the social scientist*. 2. ed.: McGraw-Hill Education (UK), 2005.
- GLASSERMAN, Paul. *Monte Carlo Methods in Financial Engineering*. New York: Springer, 2003. v. 53. (Applications of Mathematics).
- GOERTZ, Gary; MAHONEY, James. *A tale of two cultures: Qualitative and quantitative research in the social sciences*. Princeton: Princeton University Press, 2012.
- GRANSCHKE, Bruno. The Art of Staging Simulations: Mise-en-scène, Social Impact, and Simulation Literacy. In: *The Science and Art of Simulation I*. Springer, 2017. p. 33–50.

- GRECO, John. Post-Gettier epistemology. *Veritas*, Porto Alegre, v. 60, n. 3, p. 421–437, 2015.
- GRÜNE-YANOFF, Till; WEIRICH, Paul. The philosophy and epistemology of simulation: A review. *Simulation & Gaming*, SAGE Publications, v. 41, n. 1, p. 20–50, 2010.
- GUMERMAN, GJ; DEAN, J. Artificial Anasazi. *Discovering Archaeology*, v. 2, p. 44–51, 2000.
- HALAS, Juraj. Marxian “Abstraction” and Contemporary Philosophy of Science. *Epistemology & Philosophy of Science*, v. 55, n. 3, p. 122–136, 2018.
- HALLEVY, Gabriel. The Criminal Liability of Artificial Intelligence Entities-From Science Fiction to Legal Social Control. *Akron Intell. Prop. J.*, HeinOnline, v. 4, p. 171, 2010.
- HASSE, Hans; LENHARD, Johannes. Boon and bane: On the role of adjustable parameters in simulation models. In: *Mathematics as a Tool*. Springer, 2017. p. 93–115.
- HEALY, Kieran. Fuck nuance. *Sociological Theory*, SAGE Publications Sage CA, Los Angeles, CA, v. 35, n. 2, p. 118–127, 2017.
- HUMPHREYS, Paul. *Extending ourselves*: Computational science, empiricism, and scientific method. New York: Oxford University Press, 2004.
- JEBEILE, Julie. Idealizations in Empirical Modeling. In: *Mathematics as a Tool*. Springer, 2017. p. 213–232.
- JENSEN, Arthur. *Bias in mental testing*. New York: The Free Press, 1980.
- KELSEN, Hans. *Pure theory of law*. Tradução: Max Knight. Berkeley: University of California Press, 1967.
- KING, Gary; KEOHANE, Robert O; VERBA, Sidney. *Designing social inquiry*: Scientific inference in qualitative research. Princeton: Princeton University Press, 1994.
- KUHN, Thomas S. *The structure of scientific revolutions*. 50th Anniversary. Oxford: Oxford University Press, 2012.
- LENHARD, Johannes. Thought experiments and simulation experiments: Exploring hypothetical worlds. In: STUART, Michael T; FEHIGE, Yiftach; BROWN, James Robert (Ed.). *The Routledge Companion to Thought Experiments*. New York: Routledge, 2018.
- LENHARD, Johannes; KÜPPERS, Günter; SHINN, Terry. *Simulation*: Pragmatic constructions of reality. Springer Science & Business Media, 2007. v. 25.
- LIEDER, Michael; ASIF, Farazee MA; RASHID, Amir. Towards Circular Economy implementation: an agent-based simulation approach for business model changes. *Autonomous Agents and Multi-Agent Systems*, Springer, v. 31, n. 6, p. 1377–1402, 2017.
- LOPES, Marcelo Frullani; ROBERTO, Enrico. Quando um carro autônomo atropela alguém, quem responde? *El País Brasil*, abr. 2018.

- MACKENZIE, Donald; SPEARS, Taylor. 'A device for being able to book P&L': The Organizational Embedding of the Gaussian Copula. *Social Studies of Science*, Sage Publications, London, v. 44, n. 3, p. 418–440, 2014a.
- _____. 'The formula that killed Wall Street': The Gaussian copula and modelling practices in investment banking. *Social Studies of Science*, Sage Publications, v. 44, n. 3, p. 393–417, 2014b.
- MAHAJAN, Gurpreet. *Explanation and understanding in the human sciences*. 3. ed. New Delhi: Oxford University Press, 2011.
- MCCLOSKEY, Deirdre N. *The rhetoric of economics*. 2. ed. Madison: University of Wisconsin Press, 1998.
- MEDEIROS, Stanley Kreiter Bezerra. *Um estudo lógico e epistemológico do fecho epistêmico*. 2013. Tese (Doutorado em Filosofia) – Universidade Federal da Paraíba.
- MILLER, John H; PAGE, Scott E. *Complex adaptive systems*. Princeton: Princeton University Press, 2007.
- MIRAMONTES, Pedro. *Un modelo de autómatas celulares para la evolución de los ácidos nucleicos*. 1992. Tese (Doutorado em Matemática) – UNAM, Cidade do México.
- MOHIUDDIN, Syed et al. Patient flow within UK emergency departments: a systematic review of the use of computer simulation modelling methods. *BMJ open*, British Medical Journal Publishing Group, v. 7, n. 5, e015007, 2017.
- MONTEIRO BRITO JÚNIOR, Valdenor. A filosofia analítica do Direito é etnograficamente limitada? *Revista Direito GV*, v. 14, n. 1, p. 49–78, 2018.
- MOROZOV, Evgeny. *The net delusion: The dark side of Internet freedom*. New York: PublicAffairs, 2012.
- MORRISON, Margaret. Models, measurement and computer simulation: the changing face of experimentation. *Philosophical Studies*, Springer, v. 143, n. 1, p. 33–57, 2009.
- MULDOON, Ryan. Robust simulations. *Philosophy of Science*, The University of Chicago Press, v. 74, n. 5, p. 873–883, 2007.
- NGO, The An; SEE, Linda. Calibration and validation of agent-based models of land cover change. In: *Agent-based models of geographical systems*. Springer, 2012. p. 181–197.
- NONNENMACHER, Vinicius et al. Modeling and visualization individual and collective opinions towards extremism in a society. *Procedia Computer Science*, Elsevier, v. 29, p. 2412–2421, 2014.
- NOZICK, Robert. *Philosophical explanations*. Cambridge: Belknap Press, 1981.
- OBERAUER, Klaus et al. What limits working memory capacity? *Psychological Bulletin*, American Psychological Association, v. 142, n. 7, p. 758, 2016.

- OKASHA, Samir. *Philosophy of Science: A Very Short Introduction*. Oxford: Oxford University Press, 2016. (Very Short Introductions).
- PAIVA, Cibele Nogueira. *A relevância do fator humano na simulação computacional*. 2010. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Itajubá, Itajubá.
- PINSKY, Mark; KARLIN, Samuel. *An introduction to stochastic modeling*. 4. ed. Burlington: Academic Press, 2010.
- PROVOST, Foster; FAWCETT, Tom. *Data Science for Business: What you need to know about data mining and data-analytic thinking*. O'Reilly, 2013.
- RESCH, Michael M. On the Missing Coherent Theory of Simulation. In: *The Science and Art of Simulation I*. Springer, 2017. p. 23–32.
- ROEMER, John. *A general theory of exploitation and class*. Cambridge, MA: Harvard University Press, 1982.
- ROSENBERGER, Robert; VERBEEK, Peter-Paul. A field guide to postphenomenology. In: *Postphenomenological investigations: Essays on human-technology relations*. Lanham, MD: Lexington Books/Rowman Littlefield Press, 2015. p. 9–42.
- RUSSELL, STUART J.; NORVIG, Peter. *Artificial Intelligence - A Modern Approach*. 3. ed.: Pearson Education, 2010.
- RYDER, William H. A System Dynamics View of the Phillips Machine. In: PROCEEDINGS of the 27th International Conference of the System Dynamics Society. 2009.
- SAAM, Nicole J. Understanding Social Science Simulations: Distinguishing Two Categories of Simulations. In: *The Science and Art of Simulation I*. Springer, 2017a. p. 67–84.
- _____. What is a Computer Simulation? A Review of a Passionate Debate. *Journal for General Philosophy of Science*, Springer, v. 48, n. 2, p. 293–309, 2017b.
- SARGENT, Robert G. Verification and validation of simulation models. *Journal of simulation*, Taylor & Francis, v. 7, n. 1, p. 12–24, 2013.
- SCHELLING, Thomas C. *On letting a computer help with the work*. John F. Kennedy School of Government, Harvard University: Public Policy Program, 1972.
- SCHÖNEMANN, Peter H. Measurement: The reasonable ineffectiveness of mathematics in the social sciences. *Trends and perspectives in empirical social research*, Walter de Gruyter Berlin, Germany, p. 149–160, 1994.
- SIEGFRIED, Robert. *Modeling and simulation of complex systems: A framework for efficient agent-based modeling and simulation*. Springer, 2014.
- SOMMERVILLE, Ian. *Software Engineering*. 9. ed. London: Pearson Education Ltd, 2011.
- SORENSEN, ROY A. *Thought Experiments*. New York: Oxford University Press, 1992.

- SOSA, Ernest. *Epistemology*. Princeton: Princeton University Press, 2017.
- STANLEY, Jay. Pitfalls of Artificial Intelligence Decisionmaking Highlighted In Idaho ACLU Case. *Free Future*, ACLU, jun. 2017.
- STEELS, Luc. Agent-based models for the emergence and evolution of grammar. *Phil. Trans. R. Soc. B*, The Royal Society, v. 371, n. 1701, p. 20150447, 2016.
- THAGARD, Paul. *Computational philosophy of science*. Cambridge: MIT Press, 1993.
- TURING, Alan M. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London mathematical society*, Wiley Online Library, v. 2, n. 1, p. 230–265, 1937.
- VAN DER GRIENT, Leon. *On the Epistemology of Computer Simulations*. 2011. Monografia (Bacharelado em Filosofia) – Utrecht University, Utrecht.
- VAN FRAASSEN, Bas C. *The scientific image*. New York: Oxford University Press, 1980. (Clarendon Library of Logic and Philosophy).
- VENTURINI, Tommaso; JENSEN, Pablo; LATOUR, Bruno. Fill in the gap: A new alliance for social and natural sciences. *Journal of Artificial Societies and Social Simulation*, v. 18, n. 2, p. 18–29, 2015.
- VON HAYEK, Friedrich August. The pretence of knowledge. *The Swedish Journal of Economics*, JSTOR, v. 77, n. 4, p. 433–442, 1975.
- VON WRIGHT, Georg Henrik. *Explanation and understanding*. Ithaca: Cornell University Press, 2004.
- WEISBERG, Michael. *Simulation and Similarity: Using Models to Understand the World*. New York: Oxford University Press, 2013.
- WHITE, Hayden. *Meta-história: a imaginação poética do século XIX*. Tradução: José Laurênio de Melo. 2. ed. São Paulo: Edusp, 2008.
- WINSBERG, Eric. A tale of two methods. *Synthese*, Springer, v. 169, n. 3, p. 575–592, 2009.
- _____. *Science in the age of computer simulation*. Chicago: University of Chicago Press, 2010.
- WOLFF, Robert Paul. *An Introduction to the Use of Formal Methods in Political Philosophy*. Wellington, NZ: Society for Philosophy & Culture, 2013.
- YARKONI, Tal; WESTFALL, Jacob. Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, SAGE Publications, v. 12, n. 6, p. 1100–1122, 2017.
- ZAFFARONI, Eugenio Raúl et al. *Teoria do delito: introdução histórica e metodológica, ação e tipicidade*. Rio de Janeiro: Revan, 2010. (Direito Penal Brasileiro, II.1).